Transcription

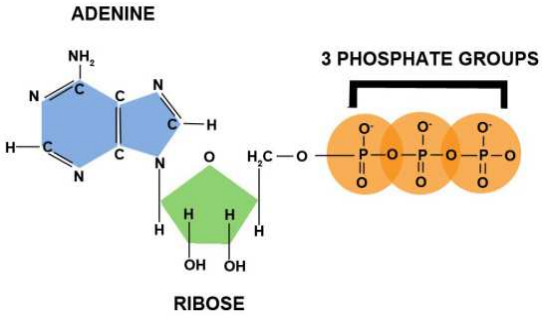
Dr.H.B.Mahesha, Yuvaraja's College, University of Mysore, Mysuru.

It is the process of transcribing or making a copy of Genetic information stored in a DNA strand into a Complementary strand of RNA (messenger RNA or mRNA) with the aid of RNA polymerase. Or The process by which RNA molecules are initiated, elongated, and terminated is called **transcription.**

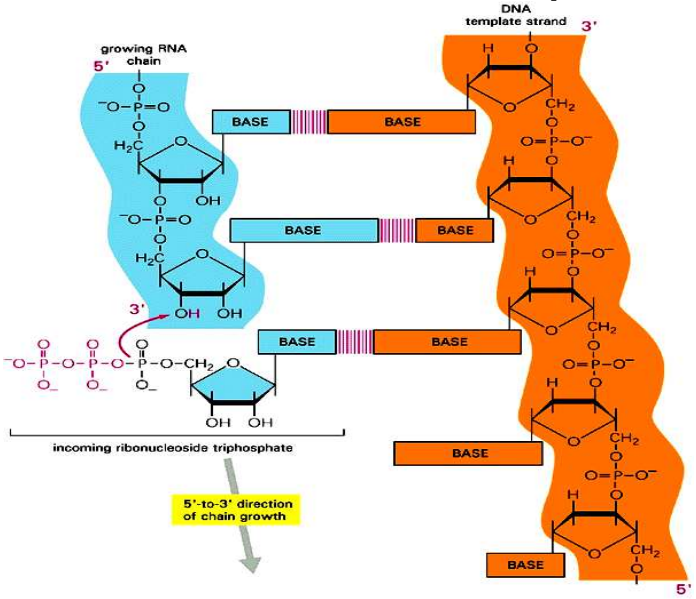
Basic Features of RNA Synthesis

The essential chemical characteristics of the synthesis ofRNA are the following

1. The precursors in the synthesis of RNA are the four ribonucleosie 5'-triphosphates (NTP) ATP, GTP, CTP, and UTP. On the ribos portion of each NTP there are two OH groups-one each on the 2'- and 3'-carbon atoms



2. In the polymerization reaction a 3'-OH group of one nucleotide reacts with the 5'-triphosphate of a second nucleotide; pyrophephate is removed and a phospho diester bond results. This is the same reaction that occurs in the synthesis of DNA.



3. The sequence of bases in an RNA molecule is determined by the base sequence of the DNA. Each base added to the growing end of the RNA chain is chosen by its ability to base-pair with the DNA strand used as a template; thus, the bases C, T, G, and A in a DNA strand cause G, A, C, and U, respectively, to appear in the newly synthesized RNA molecule.
4. The DNA molecule being transcribed is double-stranded, yet in any particular region only one strand serves as a template.

5. The RNA chain grows in the 5'~3' direction: that is, nucleotides are added only to the 3' -OH end of the growing chain-this direrion of chain growth is the same as that in DNA synthesis.
6. RNA polymerases, in contrast with DNA polymerases, are able to initiate chain growth; that is, *no primer is needed.*
7. Only ribonucleoside 5' -triphosphates participate in RNA synthesis and the first base to be laid down in the initiation event is a triphosphate. Its 3' -OH group is the point of attachment of the subsequent nucleotide. Thus, the 5' end of a growing RNA molecule terminates with a triphosphate (Figure 12-3).

The overall polymerization reaction may be written as

$$nNTP + XTP \xrightarrow[Mg^{2+}]{\text{DNA, RNA-P}} (NMP)n - XTP + nPPi$$

in which XTP represents the first nucleotide at the 5' terminus of the RNA chain, NMP is a mononucleotide in the RNA chain, RNA-P is RNA poly. merase, and PP_i is the pyrophosphate released each time a nucleotide is added to the growing chain. The Mg^{2+} ion is required for all nucleic acid polymerization reactions.

Requirements:

1. Template :- Double stranded DNA
2. Activated precursors :- ATP, GTP, UTP, CTP
3. Divalent Metal Ions :- mg^{++}, mn^{++}
4. RNA polymerase :- Enzyme

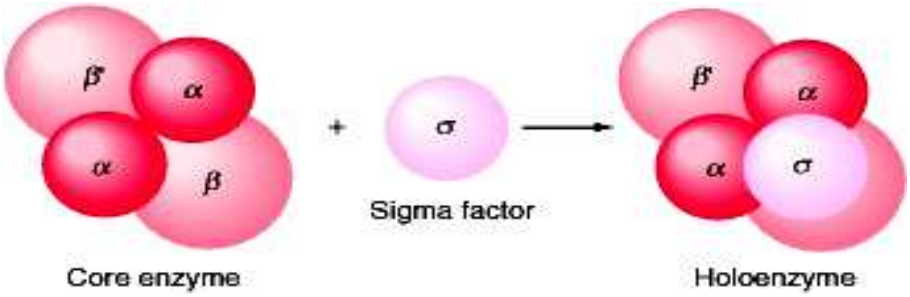
***E. coli* RNA Polymerase:**

The step-by-step elongation of the RNA chain on a DNA template is catalysed by the enzyme RNA polymerase. RNA polymerase was isolated from *E. coli* by Chamberlin and Berg (1962).

In prokaryotes a single RNA polymerase enzyme controls the synthesis of all the different types of cellular RNA. In *E. coli,* for example, the same RNA polymerase catalyses the synthesis of mRNA, tRNA and rRNA.

The entire RNA polymerase enzyme *(holozyme)* consists of a *core enzyme* and a *sigma (σ) factor*. In *Escherchia coli* the core enzyme consists of four polypeptide chains, *β, β′, α* and α*,* and is designated as , α_2, β, β'. Further details are given in the following table.

Sl. No.	Sub Unit	M. W.	Function
1	β’	1,60,000	Binding of RNA polymerase to DNA *i.e.,* Template strand
2	β	1,50,000	Helps in binding with incoming nucleotides & forms 1st phosphodiester bond formation
3	2 α	40,000	Needed for assembly of core enzyme & Promoter Recognition
4	σ	90,000	Recognition of start signal & directs to bind to the promoter region of DNA



The term **core enzyme** is used to describe the σ-free unit-namely, α_2, β, β'. The complete enzyme, α_2, β, β' σ is called the **holoenzyme.**

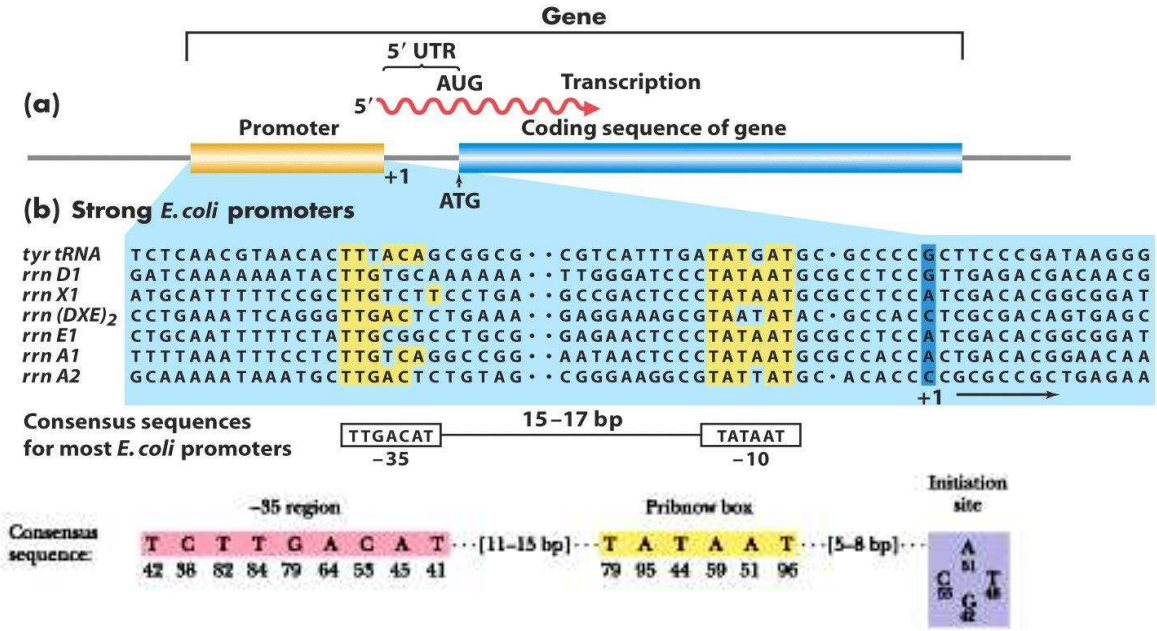
The synthesis of RNA consists of four discrete stages:

(1) binding of RNA polymerase to a template at a specific site
(2) initiation
(3) chain elongation
(4) Chain termination and release.

Several events must occur at a promoter before the initiation *i.e.,* RNA polymerase must recognize a specific DNA sequence, attach in a proper conformation, local open the DNA strands in order to gain access to the bases to be copied, and then initiate synthesis. These events are guided by the base sequence of DNA, the polymerase σ subunit (without which the promoter is not recognized) and, for some promoters, also by auxiliary proteins. The details these events are not known, but the process can be broken down into three parts-(i) template binding at a *polymerase* recognition site, (ii) binding to an initiation site, and (iii) establishment of an open-promoter complex.

Promoter : A site in a DNA molecule at which RNA polymerase and transcription factors bind to initiate transcription of mRNA.

Figure shows portions of several promoter sequences in *E. coli* is recognized by *E. coli* RNA polymer and their important features. In a region 5-10 bases to the left of the first base copied into mRNA is the right end of a sequence called the Pribnow box. A basic sequence derived from a large set of observed similar sequences is called a consensus sequence; it is obtained by comparing a large number of sequences from a particular region. If the sequences were totally unrelated, one would expect each base to occur at each position 25 percent of the time. Examination of more than 100 E. *coli* promoters has shown that the frequency of occurrence of the bases is the following, in which the subscript is the frequency:



Furthermore, there are few Pribnow boxes that differ from the consensus by more than two bases, and most that differ do so by only one base.

The Pribnow box is thought to orient RNA polymerase, such that synthesis proceeds from left to right, and to be the region at which the double helix opens to form the open-promoter complex.

Before enough sequences were known that the Pribnow box was recognizable, the first base transcribed was chosen as a reference point and numbered + 1. The direction of transcription was called downstream; all upstream bases, which are not transcribed, were given negative numbers starting from the reference. The Pribnow box is enclosed between - 21 and -4, depending on the particular promoter.

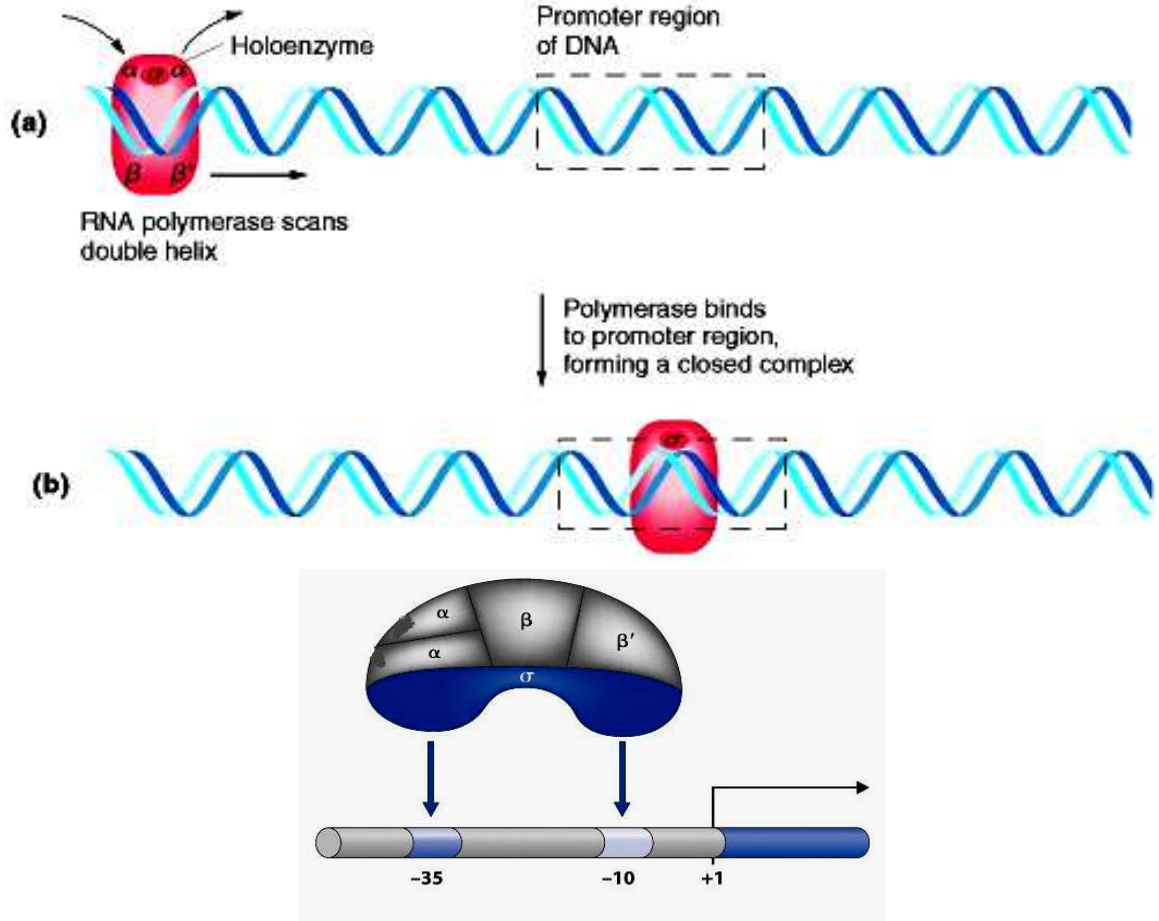
Examination of the complete sequence of the region protected by RNA polymerase indicates that for many (but not all) promoters, there is a second Important region, to the left of the Pribnow box, whose sequences in different promoters have common features (above Figure). This six-base sequence, which is called the -35 sequence and has the consensus TTGACA. may be the initial site of binding of the enzyme, when the sequence is present.

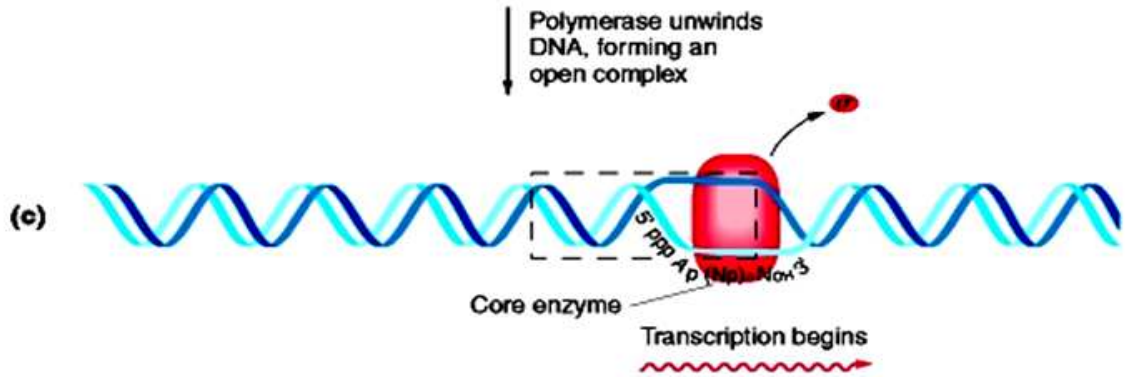
The earliest events that take place when RN A polymerase binds to a promoter and the precise roles of the - 35 sequence and the Pribnow box are not completely clear.

RNA Chain Initiation

Once the open-promoter complex has formed, RNA polymerase is ready to initiate synthesis. RN A polymerase contains two nucleotide binding sites oiled the initiation site and the elongation site, The initiation site primarily binds purine triphosphates, ATP and GTP, and ATP is usually the first nucleotide in the chain. Thus, the first DNA base that is transcribed is usually thymine. The initiating nucleoside triphosphate

binds to the enzyme in the open-promoter complex and forms a hydrogen bond with the complementary ON A base. The elongation site (also called the catalytic site) is then filled with a nucleoside triphosphate that is selected by its ability to hydrogen-bond with the next base in the DNA strand. The two nucleotides arc then joined together, the first base is released from the initiation site, and initiation is complete. In some way, the details of which are not understood, the RNA polymerase and the template strand move relative to each other, so the two binding sites and the catalytic site are shifted by exactly one nucleotide. The dinucleotide remains hydrogen-bonded to the DNA. The initiation phase is not yet complete since polymerization of the first 6-10 nucleotides is different from the process that occurs henceforth. The main difference is that the strict one-by-one-base template reading (called processivity) is not yet established. Evidence is that during this initial period, short oligonucleotides are often released from the DNA indicating that RNA polymerase stops and restarts at the initiation site. It is thought that this is a result of RNA polymerase remaining anchored to some upstream site. In an unknown way, RNA polymerase finally becomes locked in to forward motion and the elongation phase begins.





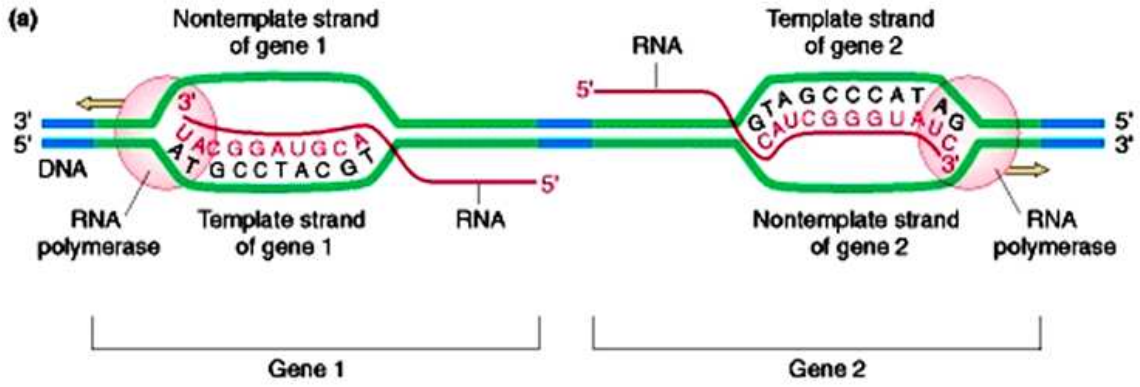
How selection of the first base copied occurs is not known. It is always within six to nine bases of the right (3') end of the Pribnow box, but neither the precise position nor the identity of adjacent bases shows any obvious pattern.

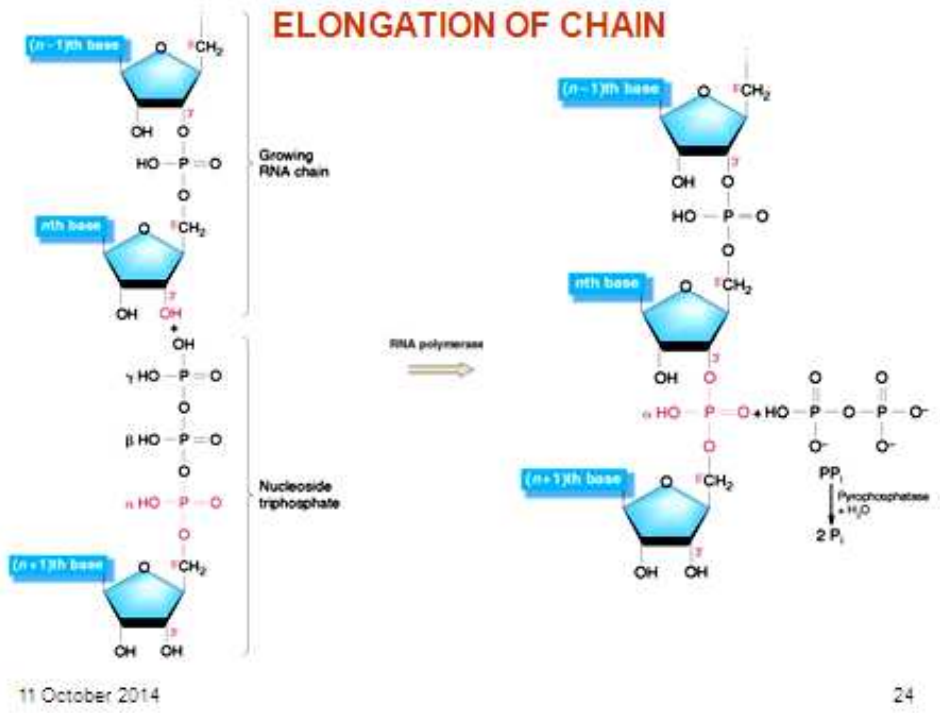
Chain Elongation

After several nucleotides (eight is the best estimate) are added to the growing chain, RNA polymerase undergoes a' conformational change and loses the *a* subunit. This marks the transition from the stuttering of the initiation phase, just described, to the stable forward movement of the elongation phase. Thus, *most elongation is carried out by the core enzyme*.

The core enzyme moves along the DNA, binding a nucleoside triphosphate that can pair with the next DNA base and opening the DNA helix as it moves. The DNA helix recloses as synthesis proceeds. The newly synthesized RNA is released from *its* hydrogen bonds with the DNA as the helix re-forms. Roughly 12 RNA bases are paired to the DNA in the open region.

A peculiarity of the chain elongation reaction is that it does not occur at a constant rate; that is, synthesis markedly slows down when particular regions of DNA are passed, then continues at the normal rate, slows down again, accelerates again, and so forth. This reduction in rate is called a **pause.** Analysis of pausing along stretches of DNA of known sequence shows that pausing frequently follows sequences that form hairpins in the RNA, but at least half of the pause sites have no recognizable features. One very recent report suggests that pausing in non hairpin regions may be associated with an increase in rate of phosphoester cleavage by RNA polymerase. Since enzymes only alter the rate and not the equilibrium of a reaction, RNA polymerase catalyzes both polymerization and degradation, with polymerization predominating vastly.





Termination and Release of Newly **Synthesized RNA**

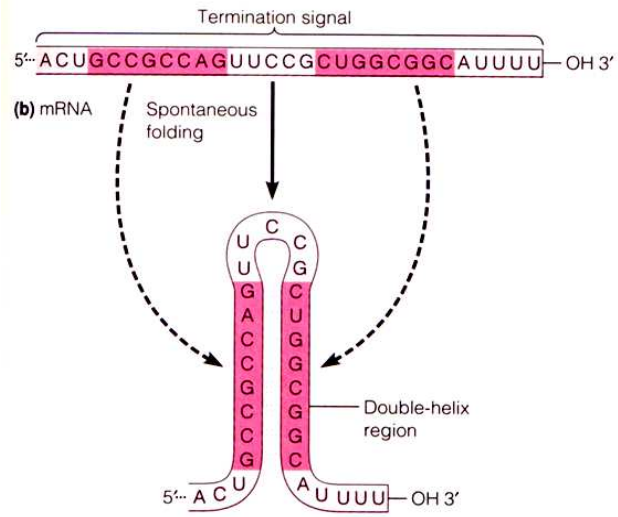
Termination of RNA synthesis occurs at specific base sequences within the DNA molecule. These sequences are of two types, simple terminators and those that require auxiliary termination factors.

Simple termination: About 100 factor independent termination sequences have been etermined in bacteria and each has the characteristics shown in Figure. There are three important regions.

1. There is an inverted repeat containing a central nonrepeating segment; that is, the sequence in one DNA strand would read like ABCDE-XYZ-E'D'C'B' A'. Thus, this sequence is capable of intrastrand base pairing, forming a stern-and-loop in the RNA transcript and possibly a cruciform structure in the DNA strands. It is possible though not yet proved that the stem-and-loop serves a purpose independent of its role in termination-namely, to render the newly synthesized RNA resistant to degradation by RNase II, an intracellular ribonuclease that is inactive against double-stranded RNA.
2. Near the loop end of the putative stem (sometimes totally within the stem) there is a high-G + C sequence. RNA polymerase usually slows down when synthesizing the corresponding RNA segment.
3. This is followed by a sequence of A·T pairs (which may begin in the putative stem) with A in the template strand, yielding, in the RNA, a sequence of 6-8 U's. The importance of this sequence is made clear by introducing deletions in the DNA that remove some of the A·T pairs in this segment. RNA polymerase still pauses in the high G + C region, but termination does not occur.

Transcription does not terminate at a unique site: in a collection of RNA molecules terminating at one site, some end with five U's and some with six U's.

How these features lead to termination is not known. It is currently believed that termination occurs when an RNA hairpin that causes a pause is formed. Evidence supporting this hypothesis is that mutations in termination sequences that remove a pause also eliminate termination.



Rho dependent termination:

'Rho-dependent termination sequences have a different form from the, Rho-independent sequences: the region in which termination occurs is preceded by a large stern-and-loop that is followed by a long tract (70-80 nucleotides) free of double-stranded segments. It is believed that the large stem-and-loop may serve to ensure that the region that follows is free of any secondary structure. The unpaired segment seems to require one or more short G + C-rich regions upstream from the termination site. A poly(U) segment is not present. Rho is a protein that binds tightly to RNA. When bound to segments rich in C (especially repeating C's), it acquires a powerful ATP-cleaving activity that is essential to its action in termination. It does not bind to either DNA or to RNA polymerase *in vitro;* however, since certain mutations in the RNA polymerase 13 subunit eliminate Rho-dependent termination and other mutations in the 13 subunit compensate for mutations in the *rho* gene that eliminate Rho activity, it seems likely that Rho and RNA polymerase interact *in vivo* in some way. The mechanism of Rho- mediated termination is not known with any certainty. The current hypothesis is that as polymerization proceeds and C-rich segments of RNA are made, the ATPase activity of Rho increases until no further polymerization is possible, because nucleoside triphosphates cannot reach RNA polymerase without being degraded. However, this explanation must' be incomplete, because certain features of Rho-dependent termination vary with the particular terminator. For example, with some sequences synthesis stops without Rho but Rho is required to prevent chain growth from continuing after a delay. At one terminator, synthesis stops completely but Rho is needed to release the RNA. Clearly, Rho has a: multifaceted activity.

The final step in termination is dissociation of both the core enzyme and the RNA from the DNA, a poorly understood step. Possibly, and this is strictly speculative, when RNA polymerase core enzyme no longer advances, the reverse reaction (degradation of phosphoester bonds) predominates, degrading the RNA right through the DNA-RNA hybrid region. Since without the σ subunit, the core enzyme cannot re-start transcription, the core enzyme also leaves the DNA.

Following dissociation of the core enzyme and the DNA, the core enzyme interacts with a free σ subunit to re-form the holoenzyme. Thus, transcription includes a σ cycle-σ falls off the holoenzyme when the elongation phase begins and rejoins the core enzyme after the core enzyme falls off the DNA (Figure).