Chapter 6. Weaver, 2/e. Mol Biol X107A.

Ch 6. The transcription apparatus of prokaryotes.

Reading notes. Also read (required):

- Gel electrophoresis, p 92 ff, especially part on proteins (for Sect C). Note that this section was assigned for Ch 2, in the context of DNA analysis. The basic ideas of gel electrophoresis are the same for both DNA and protein, but technical details differ.
- Reporter genes, p 121 (for Sect P);
- Filter binding, p 122 (for Sect D);
- DNase footprinting, p 125 (for Sect F).

Clark & Russell. Ch 6.

A. Introduction

Reminder… You may want to review the transcription section of Ch 3 in Weaver, to serve as an overview of this chapter.

This chapter discusses the mechanism of transcription. It emphasizes the initiation step in bacterial transcription; in so doing it presents the role of sigma (σ) factors in gene regulation. Initiation is the most efficient place to regulate, and is thought to be the major site of regulation.

There is much more on transcription and gene regulation in the rest of Part III along with Parts IV and V -- and even VI. But this chapter is an excellent way to start the story.

We will focus on bacterial transcription here. Although the central ideas are very similar with eukaryotes, the details, especially for initiation, are quite different. We will look at this briefly later (Ch 10-13, highlights).

Highlights

- Transcription -- an overview.
- RNA polymerase. General nature.
- The stages of transcription.
  - Initiation: major topic, separate.
  - Elongation: general biochemistry of the reaction; the DNA-RNA hybrid; how the polymerase moves; pausing and backing up. Is there proofreading?
  - Termination
    - Bacterial RNA Pol terminates at (fairly) specific sites.
    - Role of hairpin sequences and accessory proteins in transcription termination.
    - ρ-independent (intrinsic) and ρ-dependent termination.
  - "States" of the polymerase.
- Emphasis on key step of initiation. The template recognition step is especially important because of its key role in gene function, and therefore in regulation. Initiation in the broad sense also includes how the polymerase actually starts transcribing once it has found a gene. Closed and open complexes; promoter clearance.
- RNA Pol accessory factors for initiation: σ factors.

- Promoters, promoter elements; consensus sequences; role of DNA microstructure.
- Footprinting (a general method for studying DNA-binding proteins).
- Multiple σ factors, regulatory implications.
- Effect of DNA transcription on supercoiling -- and vice versa.

Comment

As we get into the core material, starting with Ch 6, it may be worthwhile to talk about the level of understanding that we expect. It is reasonable to suggest that we can discuss the material at three levels:

1. The conclusions. For example, in this Ch 6, we describe how RNA is made.
2. Experimental methods that are useful in studying such matters.
3. The historical development of how we learned particular conclusions.

Level 1 is a basic minimum. For some of you, not working in the field, it may be what you would prefer to focus on. However, for an upper division course, we really should go beyond that. Molecular biology has long been driven by the development of methods. In general, we want to develop some significant understanding of molecular biology methodology. On the other hand, this is a "principles" course, not a methods course. So we want to achieve some balance. My general approach will be to present some basic methodology, focusing on the logic of the experiments and on how the results are interpreted. However, I will not try to create a detailed history story. It is fun to do so, but those of you who are in science know that any textbook presentation of history is oversimplified.

Consistent with the above, tests will generally take it for granted that you know the basic conclusions. Few questions will solely address the conclusions. Many questions will deal with experimental methods in general, including interpretation of data. However, I do not expect you to remember history. For example, I would not ask you to remember the history of how a particular point was established, but I might present you with some data similar to that shown in the book, and ask for your interpretation.

Weaver has a strong emphasis on describing the experimental basis of the work he presents. This is good, but I would caution that the emphasis should be on following the ideas, not on remembering specific experiments. What is true is that the types of experiments that Weaver discusses often have broad applicability.

Learning molecular biology is like peeling an onion. There are layers upon layers of understanding. One first learns about the importance of DNA → RNA, then the discovery of RNA polymerase, then the discovery of the sigma (σ) factor that has a key role in initiation. Then one learns about the specific functions of σ, which are due to specific parts of the protein. Then one learns that different σ factors work differently -- and later that eukaryotes have even more complex variations. Remembering that region 2.4 of σ binds to the -10 part of

the promoter is a very specific detail. It has little meaning unless you understand all the higher level ideas. I would also suggest that this is a level of detail not worth remembering at all, because of its lack of generality.

Remember that you do not have to read and master a chapter straight through. Depending on your background, you may well want to browse the chapter, outline the key conclusions, and then go back and fill in (some of) the experimental development.

B. Common abbreviations and symbols

For "RNA polymerase", we may say "RNA Pol", or with appropriate context, just "Pol". "RNAP" is also used.

We will use several Greek letters for bacterial RNA Pol subunits, most importantly σ (sigma). The particular σ may be denoted by a superscript, either a letter, which is more or less arbitrary, or a number, which gives the approximate size in kilodaltons. σ^S and σ^{70} are examples.

Recall Ch 3 handout Sect C about use of terms such as "coding strand".

C. RNA polymerase structure

(Ch 6 Sect 1.)

Sect 1 and 2 both introduce some things that get elaborated in the rest of the chapter.

Major messages from this section

Basic subunit composition of E. coli RNA polymerase: $\alpha_2\beta\beta'\sigma$. (We usually ignore the small, mysterious and nonessential ω subunit, mentioned in Fig 6.39.) Fundamentally, this composition is of little significance, until we attach functional significance to parts.

The only functional information revealed in this section is that σ is involved in choosing specific initiation sites. It is also hinted here that σ is not always part of the enzyme. That is (as will be detailed later), σ joins the polymerase for the initiation step, then leaves, and is not part of the Pol during elongation. Thus we introduce the idea that the composition of RNA Pol varies during the transcription cycle. One might also speculate at this point that there "might" be more than one type of σ, with different specificities.

Useful methods

PAGE, Fig 6.1, but the method is introduced in an assigned part of Ch 5; see "Also read" note at top of handout. Gel electrophoresis of proteins. As usual in such physical separation methods, the migration rate of a particular type of molecule depends on the driving force

(electrical in this case), and on the shape of the molecule (via friction). In one common form of PAGE analysis of proteins, the detergent sodium dodecyl sulfate (SDS) is added. This denatures most proteins. It also coats the protein rather completely; therefore, the amount of charge is mainly due to the SDS. Considering all these effects, proteins generally separate by size in SDS-PAGE analysis.

Hybridization (Fig 6.2 and recall Ch 2). In a simple experiment, Fig 6.2a, one simply measures how much of a sample can hybridize to a reference DNA. In a competition analysis, Fig 6.2b, one tests two samples to see whether they compete with each other.

Ignore

Discussion of phage T4 life cycle and classes (phases) of RNA.

D. Promoter

(Ch 6 Sect 2.)

Major messages from this section

The main message is the identification of a specific region where RNA polymerase binds to DNA and initiates transcription. We call this the promoter.

In bacteria, a major type of promoter is characterized by regions shown in Figs 6.10 & 6.11. The -10 and -35 elements are most commonly discussed. The "optional" UP element is particularly important for extremely high levels of expression.

> Caution: The term promoter is clearly defined in a broad conceptual sense, but its precise definition is about as clear as the definition of a gene (recall Ch 3 handout Sect J). The broad concept relates to RNA Pol binding. Some specifics are well characterized for bacterial promoters, but it turns out that not all bacterial promoters follow exactly that plan. And then with eukaryotes, the concept is fine, but the details are very different (Ch 10).

Consensus sequences. Introduced here for RNA Pol and the promoter, but generalizable.

Promoters that are recognized by a particular RNA polymerase (or, rather, by a particular σ) tend to have similar sequences. In E. coli the major recognition areas for RNA polymerase (with σ^{70}) are the -35 and -10 regions ("+1" is the start site). Weaver shows the E. coli σ^{70} consensus promoter in Fig 6.10. The quality of the consensus at -10 (the Pribnow box) can be summarized by $T_{80}A_{95}T_{45}A_{60}A_{50}T_{96}$.

(The term T_{80} means that T is found at that position 80% of the time; 25% would be random.)

Consensus sequences are often derived simply from statistical analysis of what is found in Nature; that is, they describe the most common base at each site. In some cases, as in this one, the closer a specific sequence is to the consensus, the stronger the binding. Mutations that make a specific promoter closer to consensus usually make initiation more likely ("up" mutations), etc (p 140).

A third, optional promoter element, termed the UP element, is now recognized. It is upstream of the main promoter (Fig 6.11). It is needed for the high level function of the rRNA genes (among others), and is contacted by the α subunit of RNA Pol (Fig 6.30).

Lohrke et al (1999) is an example of regulation at the level of the α subunit. Bartlett et al (2000) discuss one role of FIS, shown in Fig 6.11.

Schaumburg & Tan (2000) show an additional complexity of Chlamydia promoters. Meima et al (2001) show how to isolate promoters.

Stages. Fig 6.7. See next section.

Useful method

Filter binding. See Fig 5.35 and associated text; see "Also read" note at top of handout. Note that this method is strictly empirical.

E. Stages -- of initiation, or anything

Fig 6.7 shows stages of RNA Pol binding to the promoter. The remaining sections of the Ch discuss general stages of the transcription process (initiation, etc.). It is very useful for us to divide complex processes into stages, for ease of investigation and discussion. The caution is that these distinctions are for our convenience and do not necessarily represent clear natural divisions. So be flexible when dealing with such manmade categories.

We recognize fairly distinct steps in initiation of RNA synthesis by RNA polymerase:

1. initial binding to DNA (non-specific)
2. promoter search
3. promoter recognition, which leads directly to formation of the closed complex
4. formation of the open complex
5. beginning synthesis of a chain
6. clearance of the promoter.

The third step may be considered promoter binding per se; formation of the open complex involves opening up ("melting") the DNA. We will focus on promoter recognition.

Weaver introduces the first four steps in Fig 6.7. Then Fig 6.13 overlaps with this, and continues.

F. Methods: Footprinting, protection and crosslinking experiments

The following sections involve many experiments with the general theme of trying to identify where a protein interacts with DNA. There are variations, but several involve some type of interference.

Footprinting. See Fig 5.37; see "Also read" note at top of handout.

The key idea is that the bound protein (the one we are studying) interferes with DNase digestion of the DNA. Where this interference occurs is analyzed by gel electrophoresis. The result is that some lengths don't appear, because the protein protected the DNA from being cleaved at specific sites.

> Chemical footprinting or protection. Binding of the protein to DNA may affect chemical reactions we perform on the DNA. On the other hand, specific chemical modifications of the DNA may affect protein binding. Fig 5.38 introduces one type of such experiment, and Fig 6.17 is an application. In these experiments, the only bases that get methylated by the chemical reagent (DMS) are those that have been exposed as a result of a protein binding to the DNA.

Crosslinking experiments can be thought of as a variation of chemical footprinting or protection experiments. The key difference is that the crosslinking experiment preserves the bound state of something for further analysis. (p 159 & Fig 6.37.)

> ⇒ You should be able to follow the basic logic of these experiments. However, I would not expect you to remember the specific reactions involved or the details of any particular analysis.

G. Initiation

(Sect 6.3.)

RNA polymerase recognizes a specific recognition site, called a promoter, near the beginning of a gene (or operon = group of genes transcribed together) (introduced in Sect D, above). With bacterial RNA polymerase, this base sequence recognition is due -- largely -- to a transient RNA polymerase subunit called σ.

The σ-promoted binding leads to the closed complex; this isomerizes to the open complex, and then RNA synthesis can proceed. Open complex formation involves opening up the DNA,

Figs 6.17-19. It also involves a conformational change of the core Pol itself, allowing the Pol to "hold" the DNA; Weaver introduced this conformational change back in Sect 2, Figs 6.8 & 9.

Much of Section 3 describes evidence that σ binds to the promoter, and defines the specific regions of each that are involved. Much of the early evidence was from footprinting and protection experiments (Sect F). More recently, evidence from genetic analysis and protein structures has contributed. Ultimately, we achieve the rather detailed information summarized in Fig 6.23.

Guthold et al (1999) observe RNA Pol sliding along DNA; this one-dimensional diffusion, following non-specific binding, presumably aids in finding the promoter. Naryshkin et al (2000) discuss recent work on the open complex, using crosslinking analysis.

The key property of σ is recognizing the promoter. Along the way, we learn other things about it…

σ alone cannot bind to promoters (p 148). Callaci et al (1999) and Young et al (2001) explore this.

σ reduces non-specific binding (p 151).

After RNA synthesis has started, σ dissociates from the polymerase (Fig 6.15, summarized in Fig 6.16). This weakens the strong Pol-promoter contact, and allows Pol to transcribe away from the start site. The relationship between σ release and abortive initiations (Fig 6.12) is not clear.

Weaver recognizes open complex formation (e.g., Fig 6.17) and promoter clearance (e.g., Fig 6.13), but does not try to explain them. That is fine for now; we simply note that we can recognize those steps.

Method

Rifampicin (rifamycin) is an antibiotic that specifically inhibits the initiation step of bacterial RNA synthesis (p 143). Rifampicin-resistant mutants can be easily isolated. The resistance is associated with the β subunit; you should be able to follow the logic of the mix-and-match experiment of Fig 6.32, which shows this. Later (p 163), Weaver explains that Rif blocks the exit channel for the nascent RNA. (Rifampicin is a major drug in TB treatment, but resistance to it is becoming a problem.)

H. Regulation of initiation

(End of Sect 6.3 and misc.)

The most efficient place to regulate a process is at the start. In fact, transcription is regulated at every conceivable place, but regulation of initiation does seem to be most common.

Regulation is the explicit topic of other chapters, but we note some aspects here in passing.

Anything that modulates the quality of the Pol-promoter interaction will regulate.

We have already noted that there are promoters of various strengths.

Weaver also notes briefly that there may be multiple σ's for the same cell, each with a different type of promoter specificity. More in Sect I.

Weaver notes the role of the UP promoter element and its interaction with the α subunit of Pol. Fig 6.30 has an interesting implication. In frame a, if another protein bound nearby and was able to tether α, it might also enhance transcription. In fact, it is known that some activator proteins -- proteins that stimulate transcription -- work precisely that way. Peek ahead to Fig 7.17 for an example. Also see Lohrke et al (1999).

More generally… The promoter, where RNA polymerase recognizes the gene and binds, is a good target for gene regulation. A protein tightly bound near the promoter can prevent the polymerase from binding (or functioning properly); that's the idea of a repressor. Or a protein might help the polymerase bind to the promoter, thus enhancing gene expression; that's the idea of a transcriptional activator (see previous paragraph). The role of an activator protein is particularly important if the promoter is "intrinsically" weak.

Some genes have more than one promoter, to allow for complex regulation. The promoters may use different σ factors (next section) or different activator proteins to respond to different signals. Note the title of Sect 8.4 (but don't worry about that section).

I. Sigma factors and the specificity of RNA polymerase

The σ subunit of bacterial RNA polymerase is required for initiation of RNA synthesis; σ recognizes the promoter. One can imagine having different σ factors that recognize different promoters, thus conferring different specificities on the polymerase. Weaver briefly notes this on p 147. He discusses some examples of multiple σ factors in the upcoming chapters on gene regulation.

For now, some brief comments…

Multiple σ factors were first recognized in viral systems, where the virus introduces a new σ to change the transcription specificity from host genes to viral genes.

Bacillus subtilis, a bacterium which sporulates, changes σ factors for the sporulation phase of growth. In fact, there is a cascade of σ factors during sporulation. Sect 8.3. Qiu & Helmann (2001) discuss two of the seventeen B. subtilis σ factors.

What about our old friend E. coli? For a long time we recognized only one σ factor in E. coli. However, it has become clear that there are several others, with special roles. It is likely that there are at least seven (Maeda et al, 2000). Interestingly, as information accumulates, similar types of σ factors are being found in quite diverse bacteria.

Not all σ factors work exactly the same way. As an example, Weaver introduces the B. subtilis δ protein on p 149. The details of this are not important, but there is a general idea that the various functions that are required may occur in different arrangements in different cases. σ^{54}, which transcribes the *glnA* gene, is an interesting σ variation (Ch 8 Sect 4).

More about regulation, and σ factors, in later chapters.

Jishage et al (2001) discuss an anti-sigma factor.

J. The α subunit

In Sect D, we noted how the α subunit may make DNA contacts, and may be subject to regulation. We also mentioned the regulation via α in Sect H.

In Sect 6.4, middle subsection (p 159), Weaver discusses the role of α in the overall structure of the Pol. Arguably, this might be better placed with the earlier sections on initiation, rather than with elongation. In any case, I won't pay much attention to this little subsection.

K. Perspective: Elongation and termination; states of the polymerase

> This is a new section, to try to provide some perspective for a couple of messy sections that follow (Weaver 6.4 & 6.5; L & O here). Comments and suggestions welcomed!

You may wonder why all the detailed discussion of the mechanism of elongation and termination, in sections 6.4 & 6.5. There may seem to be an excessive concern with the most intimate details of the Pol -- which amino acid is binding what. It might seem to you that elongation involves adding nucleotides, and that is done until there is a signal to stop. Why all the detail? Unfortunately, that simple view of elongation and termination begins to break down almost as soon as any attempt is made to understand how Pol moves. My purpose in this section is to give you a brief overview of what emerges from all the discussion of elongation and termination, so you will have a sense why the rest follows.

The simple view would suggest that Pol moves down the template at some constant rate -- which is in fact easy enough to measure. However, closer examination reveals that the rate is very uneven. In particular, some template sequences are easier to transcribe than others. You might be tempted to say "so what" to this. So long as the template is properly transcribed, who cares about the detailed kinetics?

The simple view would suggest -- perhaps even seem to demand -- that there are well-defined stop signals for transcription. However, closer examination reveals that is not really so. Sites characterized as stop signals increase the probability of stopping, but it is an oversimplification to suggest that a given sequence supports either elongation or termination.

As part of that, we have learned that there are proteins that modulate termination. For example, in the presence of "anti-terminator" proteins, Pol will elongate through sites that might have been stop sites.

It turns out that the elongation and termination complexities hinted at above are related. As a brief overview, we may now think of the Pol moving along the template somewhat irregularly. At its worst, it stalls -- or "pauses", as is usually said. A "paused" Pol may actually be in a different conformation than an active Pol. And we can expand that "may" of the previous sentence to suggest that there may be multiple "paused" states. Some of them may be "not serious", but some may be more serious; the Pol may even be in a state where it cannot proceed, a state some will call "arrested". A paused Pol may even back up and then "try again". Or it may terminate. In fact, it is now a reasonable view that termination is a special case of pausing, that termination can occur only from the paused state. Finally, remember those proteins that modulate termination; they may help the Pol deal with pauses.

With all this, you should see that there is interest, then, in exactly how Pol moves, and how it behaves when it has trouble moving. This is a messy field, with complicated experiments, sometimes subject to artifacts. Models for elongation and termination have changed considerably over my years of X107A; I do not know how close we are to a final truth. But the broad perspective, of trying to see how these processes occur, are related, and are modulated by proteins, is worth attention.

L. Elongation

(Ch 6 Sect 4, first and last subsections)

The basic biochemistry of elongation is straightforward. Nucleoside triphosphate (NTP) precursors donate NMP to the 3′ -OH of the growing chain (Fig 3.13).

$RNA_n + NTP \rightarrow RNA_{n+1} + PP_i$

> (RNA_n means an RNA chain of length n nucleotides. NTP is a general notation for any Nucleotide TriPhosphate. Note that we might write RNA_n as NMP_n. PP_i = pyrophosphate.)

The incoming nucleotide is activated at the 5′ end; it reacts with the 3′-OH of the growing chain. This is called growth at the 3′ end, or growth from 5′ to 3′. Recall Fig 3.13.

Pyrophosphate (PP_i) is released, and later hydrolyzed. (Mentioned in Ch 2 handout Sect F.) This PP_i hydrolysis step is important: it provides additional energy to drive the reaction toward polymerization. In fact, without PP_i hydrolysis, the energy change, ΔG, for the

polymerization reaction is about 0, making it freely reversible. There is no change in the number of phosphate bonds; the NMP is transferred from the PP_i to the growing chain. Coupling the transfer reaction with PP_i hydrolysis hydrolyzes (net) one high energy bond; thus the overall set of reactions liberates ~ 7 kcal/mole.

For initiation of an RNA chain, the first NTP is simply laid down. Thus the presence of a triphosphate on the 5′ end is a hallmark of a biosynthetic RNA end (as distinct from a cleaved RNA end or -- as we shall see later -- from a DNA end).

The work on elongation reveals that Pol subunits β and β′ are the workhorse subunits for the core functions of elongation; in fact, these are the most highly conserved subunits from one organism to another.

Several Figs, such as 6.42, show the basic transcription bubble. Weaver introduces the melting of the template DNA, and various measurements of the size of the melted region in the open complex, in Figs 6.17-19. During elongation, the melted region remains as a "transcription bubble". The size of the DNA-RNA hybrid in this bubble, long debated, is now generally accepted as about 8-9 bases. In the abstract, the size of the hybrid has no particular significance, although it may have implications for important mechanistic details. For example it is now fairly clear that the DNA-RNA hybrid is not the key contributor to maintaining the Pol bound.

Much more important is the question of what happens when Pol is stalled. While the details are still under debate, it is clear that Pol can back up, allowing the recently incorporated nucleotides to be removed, thus allowing Pol to proceed again. The argument is that the DNA-RNA hybrid senses mispairing, which may stimulate backing up and "error correction". It is debatable whether the reverse reaction provides careful proofreading, as occurs with DNA synthesis (Sect 20.2). More likely, it merely deals with gross stalls. Marr & Roberts (2000) discuss the GreA protein, an accessory factor involved in this process. The idea is that Pol backs up, thus exposing the newly synthesized region for cleavage. It is also possible that the process is coupled to template repair. Thomas et al (1998) argue that "careful proofreading" may occur.

There are three important contact areas in the transcription complex. Two of these are protein-nucleic acid: in front of and behind the growing chain end; Fig 6.36. The third is the 8-9 base RNA-DNA hybrid, in the bubble. The protein-DNA contacts maintain the basic positioning of the Pol; the hybrid region may well be sensitive to mismatches that promote backing up. Nudler (1999) reviews recent work on how Pol moves; much of the work is from his lab.

Korzheva et al (2000) use a combination of techniques to explore the structure of the transcription complex, with an attempt to understand the structural transitions that occur during elongation. Cheetham & Steitz (1999) discuss the specialized RNAP of phage T7.

M. Transcription and supercoiling

(Ch 6 Sect 4, final subsection, on Topology…, p 164)

Supercoiling of DNA relates to stress on the DNA, including unwinding (Ch 2 handout Sect H). Thus it should not be surprising that supercoiling interacts with transcription.

Transcription affects supercoiling in two ways.

During initiation, formation of the open complex opens up the DNA, this introduces positive supercoiling (if the DNA resists). Fig 6.19 makes use of this to measure the amount of DNA melting during open complex formation.

During elongation, the Pol moves along the DNA, with the DNA unwinding in front and rewinding behind the Pol. There is no net change in DNA winding (as a "constant" size bubble moves along the template), but there are local effects. The DNA ahead of the polymerase becomes overwound -- gains positive supercoils -- and the DNA behind the polymerase becomes underwound -- gains negative supercoils. Fig 6.43 lays the groundwork for how the moving Pol affects supercoiling. (I will bring a model to show this.) The effect is greatest if the DNA is anchored so that it cannot freely rotate.

You should be able to predict the direction of the effects, as stated above, from what you know about transcription and DNA structure.

Harada et al (2001) observe the DNA rotation that occurs along with transcription, if the DNA is free to move.

N. Supercoiling and transcription

(not in book)

Supercoiling and transcription can interact in another way: supercoiling can affect transcription. Initiation of transcription requires a protein to recognize a specific DNA site and requires the DNA to melt. Supercoiling affects the appearance of DNA; thus it might reasonably affect recognition events -- for better or worse. Further, the common form of supercoiling found in vivo provides a stress that helps opens up the DNA.

The most common result is that increased negative supercoiling increases promoter function. Remember that negative supercoiling is topologically equivalent to strand unwinding, and that unwinding is required for RNA polymerase to act. (Also remember that bacterial DNA typically is negatively supercoiled in vivo.)

A counter-example is the gene for the enzyme that adds negative supercoils to DNA (topoisomerase II, commonly known as DNA gyrase). Since gyrase is the enzyme that

provides the negative supercoils, it is logical that the cell interprets a deficiency in (negative) supercoiling as a signal to make more gyrase. This is seen in vivo, using mutants with varying levels of supercoiling. Further, adding a drug that inhibits gyrase, and therefore reduces supercoiling, enhances gyrase synthesis. These results describe the phenomenon, but do not make clear the mechanism.

In one case, the mechanism is clear. In vitro experiments with DNA carrying the gyrase gene provided a simple answer. DNA with varying degrees of supercoiling was used to make message. The amount of gyrase message made depended on the amount of supercoiling. Thus the effect appears to be a direct effect of the supercoiling on the efficiency of the gyrase promoter. No other proteins are involved in this in vitro system.

The bigger story here is to emphasize that the exact DNA structure, including effects of base sequence and supercoiling, affects exactly what the Pol -- especially σ -- sees. Of course, this also applies to any other protein that recognizes DNA, and to any other variable that affects the DNA microstructure. Van Komen et al (2000) show how supercoiling can affect recombination. Recall Ch 2 handout Sect I.

Fang & Wu (1998) describe an example of how the effects of transcription and supercoiling on each other may play out.

O. Termination

(Ch 6 Sect 5)

RNA synthesis stops, at more or less specific sites. In bacteria there are two general modes of termination, intrinsic and ρ-dependent. ρ (rho) is an accessory protein factor. Intrinsic termination is also referred to as ρ-independent termination.

Hairpin sequences in the RNA transcript tend to cause the RNA polymerase to pause. The hairpin (intrastrand RNA-RNA helix) tends to pull the nascent RNA chain away from the polymerase-DNA complex. Hairpins come in a variety of strengths, and this is surely a place where RNA synthesis is regulated. (Recall the cruciform structure for DNA, Ch 2 handout Sect G. The hairpin is the SS equivalent, p 166.)

The stories of pausing and then of termination relate to the nature of the elongation complex, Sect L. Active (doing elongation), paused or arrested, and terminating may all be "states" of RNA Pol. Recall Sect K, above, where we introduced this interrelationship.

At intrinsic (ρ-independent) termination sites (Fig 6.45), the hairpin feature is further enhanced by a poly(U) region, which follows immediately. RiboU-deoxyriboA base pairs are unusually weak (p 166). Thus the intrastrand RNA hairpin and the weak DNA-RNA hybrid that follows both serve to drive termination.

Gusarov & Nudler (1999) explore details of the intrinsic termination reaction. Davenport et al (2000) explore transcription and termination by single RNA polymerase molecules. Toulokhonov et al (2001) offer a more complex interpretation of how a pause/termination site works.

ρ-dependent termination. Here is a model for how ρ may work (Fig 6.50)…

Probable stages of ρ action:

- ρ gets on. On what? On the nascent RNA. ρ-binding sites on RNA are not well defined. However, they seem to include a large region (100 bases or so) without secondary structure (C-rich, G-poor) or other tightly bound proteins or ribosomes. That is, ρ enters onto a large section of free RNA. Bogden et al (1999) explore how ρ finds it target.

- The mechanism of ρ action may well be to pull the nascent RNA out of the Pol-RNA-DNA complex, by a mechanism reminiscent of the ATP-dependent helicases (which unwind duplex DNA; Ch 20).

- However, ρ achieves its goal only when the polymerase slows down, at a "pause" site. Once ρ enters the transcription complex, it apparently follows the polymerase -- until the polymerase falters. Then ρ keeps pulling, causing termination. This model supports the role of a hairpin, which may cause polymerase to slow and also aid ρ in releasing the RNA.

Because ρ acts in part via "paused" or "stalled" Pol complexes, it follows that anything that affects pausing or the strength of the paused complex may affect ρ-dependent termination. More about termination and antitermination in Ch 8. For example, Fig 8.21b shows how protein N from phage λ helps to form a complex that can prevent termination. This complex involves other proteins, and thus analysis of N and ρ helped to reveal the role of several proteins in the transcription complex.

> I encourage you to glance at Fig 8.21 at this point, just to see alternative transcription complexes with different termination properties. However, we probably will not cover this part of Ch 8, and you are not responsible for the content of the Fig, such as understanding the roles of the proteins.
>
> Weaver introduces the term processivity at Fig 8.21. Loosely, the term refers to the ability of the enzyme (the Pol) to continue down the chain without falling off. His Glossary entry is good. We will look at processivity of DNA Pol in Ch 20; good word!

Termination by eukaryotic RNA Pols is rather different; we will briefly note one part of the story in Ch 15.

Palangat et al (1998) discuss the role of pausing in allowing an antiterminator protein to bind during the HIV life cycle.

The purpose of RNA synthesis is to make a certain piece of RNA. Thus we expect that there are specific places to start and to stop. Actually, these places need not be too precise. Protein synthesis has its own start and stop signals (Part VI). mRNAs commonly contain untranslated sequences at both the front and back ends. Further, those RNAs that are not translated (e.g., tRNA and rRNA) are trimmed to proper size from the original transcript (Ch 16 Sect 1-2).

However, even if termination itself is less important than we might have expected, the story of termination reveals aspects of the transcription elongation complex. It is a reasonable view that the transcription apparatus has the ability to switch between elongation and termination complexes. Again, recall Sect K.

P. Reporters; measuring gene activity

Read pp 121-2, in Ch 5; see the "Also read" note at the top of the handout.

How do we measure how active a gene is? There are many ways. For example, we might measure the amount of mRNA made, or we might measure the amount of protein product made. What if the protein product is hard to measure? One approach is to use a "reporter" gene. A reporter gene has the "normal" regulatory elements, including promoter, but the protein-coding part has been replaced by the sequence for a protein that is easy to measure.

Weaver gives β-galactosidase, chloramphenicol acetyl transferase (CAT) and luciferase as examples of reporter genes (Ch 5; assigned part). We will learn more about the enzyme β-galactosidase in Ch 7; Silhavy (2000) discusses the history of *lac*-based reporters. Other popular reporter systems include alkaline phosphatase, green fluorescent protein (GFP), and a similar red fluorescent protein; Terskikh et al (2000) present a fancy variation of a fluorescent protein. Sussman (2001) discusses a variety of reporter systems.

We will use a reporter gene in the homework set, and from time to time later.

Q. Previews

In general outline, transcription in eukaryotes is about the same as in prokaryotes. However, the details are different, reflecting the greater complexity of the eukaryotic system, especially for initiation. Part IV, Ch 10-13, deals with this; we will look at those chapters briefly, in sequence.

In some cases RNA is "processed" after transcription and before use. We hinted at this above, in noting that some RNAs are trimmed to proper size (Sect O). Part V discusses these post-transcriptional modifications; we may discuss this briefly, at the end of the course.

As mentioned, regulation of transcription is a key part of the story of gene regulation. More about this in various later chapters.

R. Further reading

M Fang & H-Y Wu, A promoter relay mechanism for sequential gene activation. J Bact 180:626, 2/98. Transcription affects supercoiling, which then affects transcription of nearby genes.

M J Thomas et al, Transcriptional fidelity and proofreading by RNA polymerase II. Cell 93:627, 5/15/98. From in vitro work with a eukaryotic RNA Pol, they argue that proofreading to correct errors of base insertion may actually occur during transcription.

M Palangat et al, Transcriptional pausing at +62 of the HIV-1 nascent RNA modulates formation of the TAR RNA structure. Mol Cell 1:1033, 6/98. An example of regulated termination of RNA synthesis occurs with HIV. The Tat protein binds to the nascent RNA, to prevent termination. Here, they show that the actual secondary structure to which Tat binds (called TAR) is formed only after the RNA Pol pauses.

S Callaci et al, Core RNA polymerase from E. coli induces a major change in the domain arrangement of the σ^{70} subunit. Mol Cell 3:229, 2/99. Weaver notes that free σ does not bind to DNA. This is due to region 1 serving as an inhibitor of DNA binding, until σ is bound to the core. The present work extends that, and suggests that core binding to σ not only moves region 1, but also alters the spacing of regions 2.4 and 4.2, to make them more compatible with the spacing of the promoter elements.

A Viswanathan et al, Phenotypic change caused by transcriptional bypass of uracil in nondividing cells. Science 284:159, 4/2/99. (+ News, Bridges, p 62.) One common type of DNA damage is deamination of C, changing it to U. This damage is often repaired (Ch 20), but what if it is not? U is like T, and codes for A. Here, they show that transcription of U-containing DNA does indeed lead to A being inserted into the RNA. That is, an incorrect RNA and therefore an incorrect protein can be made from damaged DNA prior to being repaired. (The word 'bypass' in the title is misleading. The effect is caused by the U in the DNA being transcribed, rather than repaired.)

E Nudler, Transcription elongation: structural basis and mechanisms. J Mol Biol 288(1):1-12, 4/23/99. Review of recent work on how RNA Pol moves. Weaver presents some of the work. In general, they use carefully controlled in vitro reactions to explore the details of elongation steps.

C E Bogden et al, The structural basis for terminator recognition by the rho transcription termination factor. Mol Cell 3:487, 4/99. X-ray analysis of a crystal of the RNA-binding domain of ρ complexed with RNA. The analysis suggests why ρ binds pyrimidine-rich sequences, but seems not to explain why it prefers C.

I Gusarov & E Nudler, The mechanism of intrinsic transcription termination. Mol Cell 3:495, 4/99. A detailed analysis of how hairpin formation weakens the elongation complex. The A-rich template region that follows pauses the Pol enough to allow termination to win.

S M Lohrke et al, Transcriptional activation of Agrobacterium tumefaciens virulence gene promoters in Escherichia coli requires the A. tumefaciens rpoA gene, encoding the alpha subunit of RNA polymerase. J Bact 181:4533, 8/99. An example of the role of the α subunit of RNA Pol in gene regulation. The activator protein makes specific contacts with α. Therefore, the Agrobacterium activator protein requires the Agrobacterium α.

M Guthold et al, Direct observation of one-dimensional diffusion and transcription by Escherichia coli RNA polymerase. Biophys J 77: 2284-2294, 10/99. Use of scanning force microscopy to study individual molecules. Of particular note, they observe RNA Pol "sliding" along the DNA, in a process that can be considered as one-dimensional diffusion. It is likely that Pol makes contact with DNA originally via non-specific binding, then slides along the DNA and finds a promoter. One-dimensional diffusion speeds up the process of finding promoters.

G M T Cheetham & T A Steitz, Structure of a transcribing T7 RNA polymerase initiation complex. Science 286:2305, 12/17/99. The phage T7 RNA Pol is simple compared to regular Pols, mainly because it only recognizes one fairly specific promoter. Otherwise, it carries out the same basic functions. Weaver introduces this phage RNA Pol in Sect 8.2. (T7 Pol, along with its corresponding phage promoter, is often used in special gene constructs.)

H Maeda et al, Two extracytoplasmic function sigma subunits, σ^{E} and σ^{FecI}, of Escherichia coli: promoter selectivity and intracellular levels. J Bacteriol 182:1181-1184, 2/00. A brief note about the two least familiar of the seven σ factors of E coli. These σ factors typically transcribe genes for "extracellular" functions, including periplasmic and outer membrane proteins.

R J Davenport et al, Single-molecule study of transcriptional pausing and arrest by E coli RNA polymerase. Science 287:2497, 3/31/00. (+ News, Buc, p 2437.) Another example of studies of single molecules, using "optical tweezers". (Recall Allemand et al, 1998; Ch 2 handout). Previous work had established that RNA Pol is a quite powerful motor protein, exerting a force of 14 pN -- similar to or greater than that from motor proteins such as those associated with muscle or microtubules. Here they study the transcription rate of individual RNA Pol molecules; they focus on pause events. Their key findings so far are that Pols move at different rates, and have varying susceptibility to pausing.

M S Bartlett et al, Regulation of rRNA transcription is remarkably robust: FIS compensates for altered nucleoside triphosphate sensing by mutant RNA polymerases at Escherichia coli rrn P1 promoters. J Bacteriol 182:1969-1977, 4/00. Weaver notes the role of the FIS protein in activating rRNA synthesis (Fig 6.11). Another regulator of rRNA synthesis is the concentration of the initiating nucleotide. More NTP leads to more synthesis. They have mutants with an altered response to NTP level; they require a higher level of NTP for efficient transcription. Surprisingly, these mutants grow (and make rRNA) fine. Why? Apparently because the role of the activator FIS protein is increased. The work shows mechanisms of rRNA regulation, but also shows how the mechanisms interact.

N Naryshkin et al, Structural organization of the RNA polymerase-promoter open complex. Cell 101(6):601, 6/9/00. Analysis of protein-DNA interaction by crosslinking.

N Korzheva et al, A structural model of transcription elongation. Science 289:619, 7/28/00. They use a combination of x-ray structure analysis and crosslinking to explore the contacts between the nucleic acids (template and product) and the Pol during transcription. As a result, they deduce features of the elongation process.

R Wooster, Cancer classification with DNA microarrays - is less more? TIG 16(8):327, 8/00. A gene chip contains samples of DNA from thousands of genes. RNA samples prepared from various conditions/tissues/people (or whatever is being studied) are hybridized to the chip, and computer analyzed to show the pattern of gene expression. This paper discusses use of gene chips for analyzing patterns of gene expression in cancer. The idea is that such patterns may distinguish classes of cancer that are not otherwise distinguishable, and this may have implications for proper therapy. Also see Hamadeh & Afshari (2000), below.

C S Schaumburg & M Tan, A positive cis-acting DNA element is required for high-level transcription in Chlamydia. J Bacteriol 182:5167-5171, 9/00. Chlamydia promoters and RNA Pol seem similar to those in E. coli, but now they show that an AT-rich region in the spacer, near the -35 element, greatly stimulates transcription. They do not know whether this is due to contact with one or another protein, or to an effect on DNA shape. If nothing else, the work is a caution that our usual description of the promoter region is over-simplified.

S Van Komen et al, Superhelicity-driven homologous DNA pairing by yeast recombination factors Rad51 and Rad54. Molecular Cell 6(3):563, 9/00. Negative supercoiling can affect recombination in ways similar to what we discussed for transcription. Same basic ideas.

A Terskikh et al, “Fluorescent timer”: Protein that changes color with time. Science 290:1585, 11/24/00. (+ News, Chicurel, p 1478.) They start with a red fluorescent protein, and make a mutant whose color changes with time. As a result, they can measure not only the level of transcription but its timing.

H Hamadeh & C A Afshari, Gene chips and functional genomics. Amer Sci 88:508, 11/00. Good introductory article on the use of gene chips (or “arrays”) for analyzing expression levels of very large numbers of genes. For more, see the X107 web page section on Gene chips. Also see Wooster (2000), above.

T J Silhavy, Gene fusions. J Bacteriol 182:5935-5938. 11/00. Commentary. We introduced reporter genes in Sect P. Reporter genes are a modern application of gene fusions, and Silhavy tells some of the history of this broader story. Much of it makes use of the *lac* system. Nice reading.

M T Marr & J W Roberts, Function of transcription cleavage factors GreA and GreB at a regulatory pause site. Molecular Cell, 6(6):1275-85, 12/00. GreA and GreB are part of the machinery for cleaving stalled transcripts, so Pol can try again. The details of the mechanism remains unclear. This paper shows that Gre-induced cleavage is necessary for normal progression through well-identified pause sites.

Y Harada et al, Direct observation of DNA rotation during transcription by Escherichia coli RNA polymerase. Nature 409:113, 1/4/01. Use of single molecule biophysics.

Y Cheng et al, A long T•A tract in the upp initially transcribed region is required for regulation of upp expression by UTP-dependent reiterative transcription in Escherichia coli. J Bacteriol 183:221-228, 1/01. This paper illustrates a complexity of initiation; it relates to choice of start site, promoter clearance, and regulation. The key phenomenon is RNA Pol "stuttering" -- transcribing the same bases over and over again. This occurs in repeated sequences, and is presumably due to template slippage. What makes this case interesting is that the stuttering varies with conditions -- in a way that is useful. The gene in question involves pyrimidine metabolism, and -- reasonably -- its function depends on the level of pyrimidines. But the mechanism is odd. The level of UTP affects the choice of transcriptional start site. And the site chosen when [UTP] is high is in a position where stuttering is "a problem", and few transcripts get made. Thus high [UTP] serves to reduce gene function, due to altered start site selection which enhances stuttering.

J Qiu & J D Helmann, The -10 region is a key promoter specificity determinant for the Bacillus subtilis extracytoplasmic-function σ factors σ^X and σ^W. J Bacteriol 183:1921-1927, 3/01. Bacillus subtilis has long been a model system for studying σ factors, and was known to have many of them. Current count is 17, 7 of which were uncovered by the genome sequencing. Here they define a very specific recognition difference between σ^X and σ^W. Interestingly, the recognition information led them to postulate which amino acids of the σ made critical contacts, but changing these (the kind of experiment Weaver discusses on p 149) did not lead to the expected changes.

I Toulokhonov et al, Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. Science 292:730, 4/27/01. They show a "tripartite interaction" of a hairpin in nascent RNA, a particular region of the Pol that can block the RNA exit channel, and a regulatory protein that can promote termination. They interpret all this in terms of control of the state of the Pol. The major point may be their suggestion that the primary effect of hairpin formation is to cause an allosteric change in the Pol, inhibiting addition of the next nucleotide to the growing chain.

J Xu et al, Function-based selection and characterization of base-pair polymorphisms in a promoter of Escherichia coli RNA polymerase-σ^{70}. J Bacteriol 183:2866-2873, 5/01. See hw.

M Jishage et al, Mapping of the Rsd contact site on the sigma 70 subunit of Escherichia coli RNA polymerase. J Bacteriol 183:2952-2956, 5/01. Rsd is an anti-sigma factor; it is a protein that binds to σ and inhibits it. An increasing number of anti-sigmas are being found; apparently, turning σ off is an important part of gene control. Here they show that Rsd interacts with σ in region 4, and probably prevents σ from making critical contacts with core Pol and/or regulatory proteins.

R Meima et al, Promoter cloning in the radioresistant bacterium Deinococcus radiodurans. J Bacteriol 183:3169-3175, 5/01. Cloning promoters is a way to find out what promoters are like in a new organism. The basic logic is to use a vector in which a reporter gene fails to function only because it lacks a promoter. Cloned sequences that allow reporter functioning are, by definition, promoters. Cf other cloning vectors, Sect 4.3.

B A Young et al, A coiled-coil from the RNA polymerase β′ subunit allosterically induces selective nontemplate strand binding by σ^{70}. Cell 105(7):935-944, 6/29/01. They identify a particular region of the Pol β′ subunit that is responsible for allowing σ to bind the promoter.

H E Sussman, Choosing the best reporter assay. The Scientist 15(15), 7/23/01, p 25. Discussion of reporter systems, with pro and con features. Includes information on commercial availability. More extensive information is at the online version: http://www.the-scientist.com

G Bar-Nahum & E Nudler, Isolation and characterization of σ^{70}-retaining transcription elongation complexes from Escherichia coli. Cell 106(4):443-451, 8/24/01. (+ accompanying article by Mukhopadhyay et al, p 453.) Ch 6 hw #13 points you to a new article from Nudler. I finally read it (and also the accompanying one). Both papers report that σ does not always dissociate from Pol for elongation. They even suggest that the non-dissociation is useful, in making initiation of the next round more efficient. This new idea goes against the long held view of the σ cycle, which includes σ dissociation as a key part of the transition from initiation to elongation. Even if the new work is correct, I suspect that there is some functional change of the σ-Pol interaction that allows elongation; perhaps complete dissociation is not necessary, but it seems likely that some change is necessary. Time will tell!

S. Erratum

p 156, Fig 6.33, part b. In both structures, there is an extra O between the adenosine and the first phosphate.

T. Homework

Weaver provides a long and useful problem set. I would like for you to go through this, though not necessarily doing everything. (Some questions go more into experimental detail than we need, as discussed above.) We can discuss some of these as needed.

It may be good to re-read the "Comment" at the end of Sect A, above, for some perspective on what your goals are as you go through all these problems.

1. RNA is easily hydrolyzed (by mild base), between the 5′ position of the sugar and the 3′-phosphate of the adjacent nucleotide. Thus after complete hydrolysis all internal nucleotides have been converted to 3′-NMP. (3′-NMP means 3′ nucleoside monophosphate, the nucleoside with one phosphate group on the 3′ end. We might also write it as Np, where the p on the right side of the N refers to the 3′ side.) What is the 5′ nucleotide of the RNA chain converted to by this kind of hydrolysis? The 3′ nucleotide? Make a diagram using the shorthand shown in Fig 3.13 as a guide.

2. A simple statement… σ is designed to bind to promoters as strongly as possible. What's wrong with this simple view? For each issue you raise, explain how σ deals with it.

3. This question explores -- one at a time -- possible ways for supercoiling to influence the level of transcription of a gene.

a. Consider, for now, only the effect of supercoiling on the polymerase opening up the DNA. What effect would you expect increased negative supercoiling to have on transcription? Explain.

b. Now consider only the effect on the polymerase recognizing and binding to the promoter. What effect would you expect increased negative supercoiling to have on transcription? Explain.

4. Gyrase (topoisomerase II) is the enzyme that adds negative supercoils to bacterial DNA. The gyrase promoter is more active as negative supercoiling decreases (an unusual but "logical" effect). A special construction is made, with the promoter for gyrase in front of the *galK* gene, which codes for the easily assayed "reporter" enzyme galactokinase. (For the nature of reporters, see Sect P.)

a. Sketch the construction.

b. With this construction, what effect would you expect addition of the drug nalidixic acid ("Nal"), a gyrase inhibitor (which inhibits the activity of the enzyme), to have on the synthesis of galactokinase? Explain.

5. The level of supercoiling of bacterial DNA in the cell is due to the balance between the activities of gyrase (topoisomerase II), which puts in negative supercoils (and/or removes positive supercoils), and topoisomerase I, which relaxes negative supercoils. In Ch 2 we noted that DNA is commonly underwound in the cell.

Let's look at the effect of transcription on the level of supercoiling. Consider a small plasmid, carrying a highly transcribed gene. It is in a mutant host that is highly deficient in gyrase. What would happen to the supercoiling of the plasmid DNA? Explain. (Assume that the proposed effect of transcription on supercoiling plays a dominant role in determining the supercoiling. For simplicity in thinking about it, you can assume that you start with a plasmid with no supercoiling; the direction of the effect is what is interesting.)

6. Consider a piece of DS DNA, with a 12-base "bubble" (where the two strands are not complementary to each other). Next, bind a short piece of RNA which is complementary to one strand in the bubble. Finally, add RNA polymerase (and nucleotides, etc.).

a. Sketch the nucleic acid structure, with the Pol.

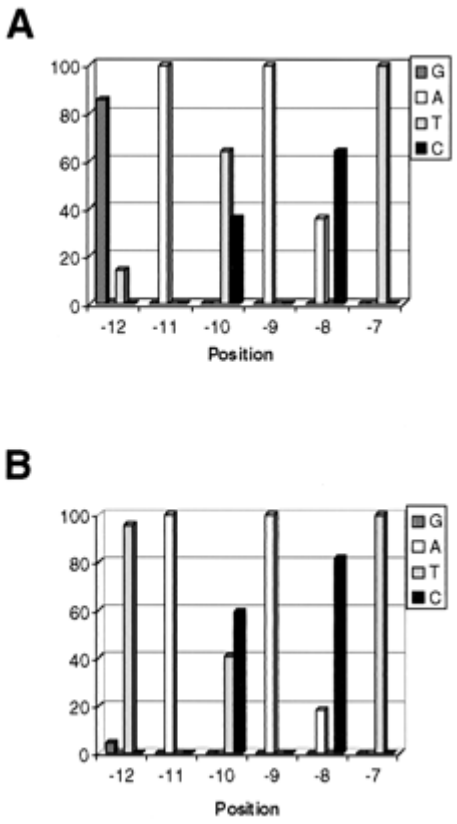
b. Would you expect that the RNA polymerase would extend the RNA chain? Explain the basis of your prediction.

7. An operational definition of the open complex is that it is resistant to adding heparin (a polyanion; mentioned p 141). In recent work, a lab reported exploration of the importance of sequences in the -10 region for formation of the open complex. In each of two parts of the experiment, they made all possible single base changes in the -10 region, and then asked which sequences worked the best. In one part of the experiment (Part B in the Figure below)

they simply tested to see which promoter sequences formed heparin-resistant complexes most rapidly. In a second part (Part A), they did the same thing, but using a very low concentration of RNA Pol.

a. What specific properties of the promoter were being selected for in these two experiments? Be sure that your level of explanation makes clear the difference between the two parts.

The following Fig shows their results.



b. What is the major difference they found between the optimum promoter sequences under these conditions?

c. At which positions are their results virtually identical for the two selections? Why do you think this is so?

d. Write out the consensus sequence they found for Part A.

e. The consensus you wrote above (d) does not agree with the -10 consensus Weaver tells you. Why? Is Weaver wrong?

f. If someone else did a similar experiment, is it plausible that they might find different results? Try to offer <u>specific</u> reasons for your answer.

8. Look at the experiments of Figs 5.37 and 6.20. What is the key difference in experimental strategy between the two experiments? Compare the information you obtain from them. Would you consider them both to be in the same general class of experiments (footprinting, discussed in Sect F, above)?

9. You have a short piece of DNA that you believe is involved with DNA replication. You insert this piece into a plasmid vector. You now carry out two restriction analyses of the new plasmid. (Restriction analysis involves treatment of the DNA with an enzyme that cuts DNA at a specific sequence. See pp 60-62 if you want a discussion, but no further knowledge is needed to do this question.) Both use the same restriction enzyme, EcoR1. In one restriction reaction, you add a protein that is required for DNA replication; this is labeled “test” below. The gel electrophoresis results for the purified DNA after restriction analyses are:

Test	Control	Size, kb
-	-	3.0
-		2.0
	-	1.5
	-	0.5

Your conclusions? Explain.

10. In Sect O, above, I explain that a hairpin and poly(A)-rich template are required for intrinsic (ρ-independent) termination. In the first full paragraph, p 168 right hand column ("Together, these results show… "), Weaver says quite to the contrary. Please resolve our little disagreement. Try to do this by reading over the section (in text and handout), and not by focusing on his summary. That is, try to develop your own understanding, from reading the evidence, rather than just reading his understanding. (In the answer section, I offer some more specific thoughts on how to approach this. But again, please give it a serious try yourself before reading that. Neither Weaver's summary nor my answer need be the way you go.)

11. Improve the chapter summary.

12. Internet question. Much of the work on elongation was done by E Nudler's group (e.g., Fig 6.37 and Nudler, 1999). A reasonable question would be to ask whether he has published more recent work. To check this, you want to do a literature search. See my Library page for how to search Medline. Go to the Medline/PubMed site, and enter Nudler as the author search term. See what you get, and perhaps browse the most recent item.

In the answer section below, I briefly note the expected output.

> Those who are experienced searchers can skip this; you are not responsible for the output. The point is to get those who are not familiar with Medline to get started. Searching on an author, with an uncommon name, is a simple search.

U. Partial answers

1. The 5′ end is converted to pppNp, the 3′ end is converted to the nucleoside.

> The 2′-OH group is key to this base-catalyzed hydrolysis (which is why it occurs with RNA but not with DNA). The first intermediate involves the P group being attached to both the 2′ and 3′ OH groups, breaking the P bond with the next nucleotide. This cyclic phosphodiester is then hydrolyzed to a mixture of 2′- and 3′-NMP. (The statement in the question that the internal nucleotides become 3′-NMP is not exactly correct. They actually become a mixture of 2′- and 3′-NMP.)

2. For example… Binding of free σ to DNA would not seem to be so useful. Solution: σ binds to DNA only when bound to Pol. (Comment. One might consider the possibility that σ selects a promoter site, then recruits Pol to it. This does not occur for bacterial σ, but this sequence of events probably does occur in eukaryotes.)

Another issue?

3. a. Negative supercoiling is topologically equivalent to unwinding the primary helix. Thus, the DNA is already partially unwound and it would be easier for polymerase to open up the negatively supercoiled DNA.

b. The polymerase makes specific contacts with specific parts of the DNA. We know that there are promoters of various strengths, which allow various quality of contact with the polymerase. One might expect that the quality of the contact in a specific case would be sensitive to supercoiling of the DNA, since that does affect its three dimensional structure. However, it is just as reasonable that the effect might be in either direction, depending on the specific case.

> The supercoiling might affect any regulatory protein that makes specific contacts with the DNA structure. Depending on whether the protein acts positively or negatively, the effect of increased binding would be to increase or decrease gene function, respectively. Further, the same point could be made for any other DNA modification, such as methylation or binding of some other protein.

The net result (from parts a & b) is that supercoiling can affect transcription by multiple mechanisms; at least some of the mechanisms can act in either direction.

The big message is the importance of the exact DNA microstructure.

4. a. _____+1_____ (where +1 marks the transcriptional start site)
p_{gyr} <-- galK gene ------>

b. Nal inhibits gyrase, thus reduces the degree of negative supercoiling. This would stimulate the gyrase promoter, and therefore the production of galactokinase, which is being made from that promoter in this case.

5. The DNA would be positively supercoiled. Transcription causes regions of positive supercoiling and negative supercoiling. Topo I would relax the negative supercoils. However, in the absence of Topo II (gyrase), the transcription-induced positive supercoils would accumulate.

> This question summarizes a key experiment that showed the effect of transcription on supercoiling. It's necessary to use a small plasmid with only one active gene, or else the effect is likely to be diluted out. Indeed, positively supercoiled DNA can be found, when Topo II is reduced either by mutation or drugs (such as Nal; see question 4). The positively supercoiled DNA is found only when the gene is being actively transcribed. (p 165)

6. The first point is to recognize that the assembled nucleic acid structure does "look like" an RNA elongation complex (e.g., Fig 6.42) -- except, of course, it lacks polymerase. So, can Pol add into an elongation complex? What do you think? Why?

> Sidorenkov et al (1998 -- in Weaver's list) used this methodology, which had been developed a few years earlier by von Hippel's lab.

7. a. One part simply selects for rate -- for fast formation of open complex. The other part selects for a high affinity of the promoter for the Pol -- because of the low concentration of Pol. In fact, they label Part A as "affinity-selected" and Part B as "rate-selected".

b. Position -12. A <u>strong</u> preference for G in Part A, vs a <u>strong</u> preference for T in Part B.

c. -7, -9, -11.

e. The work here shows that the best promoter sequence depends on the conditions (Part A vs Part B of the Figure). The consensus in Weaver is merely a statistical compilation of which bases are found most commonly. It is commonly thought that the consensus promoter is very strong. "Consensus" is a general term; one needs to make sure you know the basis of any particular statement of a consensus sequence.

f. What are some specific reasons that would result in different "best" sequences?

> This question is based on Xu et al, 2001. The Fig given is slightly modified from their Fig 2.

8. Look at the properties of the nuclease used to digest the DNA in the two experiments. If you haven't already, do that before reading further.

The nuclease of Fig 5.37 is an endonuclease, which nicks the DNA at "random" sites within the chain. It thus creates a range of fragments, of all lengths except those protected by the bound protein. The nuclease of Fig 6.20 is an exonuclease, which degrades only from the 3′ end. It degrades from that end to the block, creating one specific fragment; that fragment is defined by one end of the binding site of the protein on the DNA. Note then that the first experiment helps to define both ends of the protein binding site, whereas the second only defines one end of it.

9. First, let's explicitly state the observation: The presence of the replication protein causes two fragments to now appear as one.

It thus seems that the added protein binds to the plasmid close enough to one of the EcoR1 sites that it blocks that site from being restricted.

10. Some things to think about (you may have more)…

– In what order did we learn the various pieces of evidence?

– Did later evidence "negate" or "elaborate on" the original model?

– Which model is "higher level"?

– Which model is most relevant?

12. The output will vary depending on when you do the search. At my last check, the PubMed site gave a new article in the Aug 24, 2001, issue of Cell. Here is a brief comment about that article, which I added later to this handout:

G Bar-Nahum & E Nudler, Isolation and characterization of σ^{70}-retaining transcription elongation complexes from Escherichia coli. Cell 106(4):443-451, 8/24/01. (+ accompanying article by Mukhopadhyay et al, p 453.) Both papers report that σ does not always dissociate from Pol for elongation. They even suggest that the non-dissociation is useful, in making initiation of the next round more efficient. This new idea goes against the long held view of the σ cycle, which includes σ dissociation as a key part of the transition from initiation to elongation. Even if the new work is correct, I suspect that there is some functional change of the σ-Pol interaction that allows elongation; perhaps complete dissociation is not necessary, but it seems likely that some change is necessary. Time will tell!