Chapters 10-13. Weaver, 2/e. Mol Biol X107A.

Excerpts from:
Ch 10. Eukaryotic RNA polymerases and their promoters.
Ch 11. General transcription factors in eukaryotes.
Ch 12. Transcription activators in eukaryotes.
Ch 13. Chromatin structure and its effects on transcription.

Reminder. Sample tests for Tests 2 & 3 are at the web site, with answer keys. Current coverage of material may not correspond entirely to what is on sample tests, but should be close.

A. Introduction

Time and priorities do not permit much coverage of eukaryotic transcription in X107A. The goal here is to briefly introduce you to the highlights of eukaryotic transcription.

We will cover only limited parts of these chapters. The basic plan is as follows:

- Establish that there are three RNA polymerases in eukaryotes.
- Then focus on Pol II, the one that makes mRNA.
- In discussing Pol II, the basic story is that eukaryotic transcription is similar to prokaryotic transcription, but much more complex at the initiation step. We will establish that it is complex, but not try to master the complexity.
- We briefly note that a key protein in the Pol II system is also part of the story for the other Pols, thus showing that the three eukaryotic RNA Pols are perhaps less distinct than we once thought.
- The role of chromatin structure in eukaryotic transcription. This is a fairly new and rapidly developing area; its importance is being increasingly recognized.

Consistent with this plan, we will cover the following sections:

- Ch 10
  - Sect 1 -- briefly
  - Sect 2 -- Pol II part
  - Sect 3

- Ch 11
  - Sect 1 only, except to briefly note the role of TBP with the other Pols. In Sect 1, we will emphasize subsections on the preinitiation complex, TFIID, and the holoenzyme.

- Ch 12
  - none. We note only that much regulation is at the level of activation of initiation, and can be very complex.

- Ch 13
  - A brief overview of the entire chapter. However, the key part for us is Sect 3, starting with Pol II (p 393-end).

Even within these “assigned” sections, our coverage will be superficial. The goal is to acquaint you with eukaryotic transcription, not to detail it. Please use this handout as a guide to level of coverage.

B. Abbreviations

TF = transcription factor(s); individual TF may be named with various following numerals (to identify which Pol) and letters (to identify the specific factor). TBP = TATA-binding protein; TATA refers to a promoter element with that consensus sequence. TAF = TBP-associated factor(s); individual TAF are usually named with a number indicating the nominal size, in kilodaltons. TRF = TBP-related factor.

HAT = histone acetyl transferase or (more informally) histone acetylase; HDAC = histone deacetylase.

C. Transcription in eukaryotes

As one might imagine, the basic biochemistry of transcription is the same in eukaryotes as in prokaryotes (Ch 6). The major difference is at the initiation stage. Finding genes is much more complicated in the larger eukaryotic genome. This is because of both the larger size per se and the greater complexity of the chromatin structure. The big story is that both the promoter and RNA Pol are more complex in eukaryotes, and many more proteins are involved in putting together an effective initiation complex. This was all quite mysterious until just a few years ago, but now much is understood.

D. Multiple RNA polymerases in eukaryotes; transcription factors

(Ch 10 Sect 1)

Eukaryotic cells use three (operationally) distinct RNA polymerases to transcribe distinct classes of RNA; Fig 10.2. RNA polymerase II is responsible for making mRNA, and is therefore at the center of the gene regulation story. We will emphasize Pol II.

> RNA polymerases I and III have relatively specialized roles. Pol I makes rRNA, and Pol III makes several small RNAs such as tRNAs. (Table 10.1, p 264.)
>
> α-amanitin is a useful tool for distinguishing Pol II; Figs 10.6 & 7. (On p 273, Weaver notes that the sensitivity of Pol II to amanitin is based in the large subunit. I do not know the actual basis of amanitin action or sensitivity.)

The major subunits of all RNA polymerases (bacterial and the various Pols in eukaryotes) are quite homologous; Fig 10.11. Further, the three eukaryotic Pol share some subunits (Table 10.2, p 266). We now know that they also share a key TF (TBP; see Sect H, below).

Each Pol requires a variety of transcription proteins and DNA sites. We can group the transcription proteins into three categories: the polymerase per se, general transcription factors and specific transcription factors. We can group the DNA sites into two categories: promoters

and regulatory sites. We will examine these types of proteins and sites in the following sections (but mainly for Pol II).

E. RNA polymerase II

(also in Ch 10 Sect 1)

The complexity of the Pol II itself is introduced in the subsection starting on p 272. The major points are…

- many subunits;
- some are shared with other Pol;
- major subunits are similar to those of bacterial Pol (as expected; they do transcription).

For yeast, the genes for all 12 subunits of RNA Pol II have been cloned. Mutants in ten of these are inviable, as would be expected. Interestingly, mutants in two subunits are viable, at least under some growth conditions (p 269). Pol II is very similar in humans.

A characteristic feature of Pol II is a long C-terminal domain (or tail; CTD) on the large subunit (B1, the β′ homolog; pp 269, 272). It's a (near-perfect) repeat of a heptameric sequence (Tyr-Ser-Pro-Thr-Ser-Pro-Ser). This poly-heptamer tail is remarkably conserved among all eukaryotic Pol IIs. The CTD is essential in vivo, but not always necessary in vitro. Phosphorylation of the CTD, probably by TFIIH, is apparently part of the process of "promoter clearance", where the Pol moves away from the promoter region. More about TFIIH below.

The CTD seems to carry various proteins, and thus may be important in various phases of transcription. Calvo & Manley (2001) illustrate this.

Klug (2001) introduces the latest structural studies on Pol II; cf Fig 10.23.

F. Method: linker scanning mutagenesis

Some of the discussion of eukaryotic promoters involves work trying to find their key components (or "elements"). Work with mutants plays a role in this. Some mutants are simply deletions. One tests various deletions, and if there is an altered phenotype, one might infer that the deleted region was important. Figs 10-26 & 27 are examples of such work.

However, a criticism of using such a set of deletion mutations is that the deletion alters the spacing between the remaining sequences. Therefore, an alternative method was developed, with the goal of testing deletions without changing the spacing of the remaining sequences. The basic idea is to replace test sequences with a common "junk" sequence, rather than to delete the test sequences. The "junk" sequence is called a linker, and the method is often

called linker scanning mutagenesis. Fig 10.30 shows the idea; Figs 10-31 & 32 show some results.

G. Promoters; elements; enhancers

(Ch 10 Sect 2-3)

In bacteria we use the term "promoter", in the narrow sense, to refer to the site of binding (and/or recognition) by RNA polymerase. Thus we refer to the -35 and -10 regions (or "elements") of the common E. coli promoter (Fig 6.10). We also use the term in a broader sense: the RNA Pol binding site plus associated regulatory sites; we sometimes call this the "promoter region."

In eukaryotes, the narrow sense of the term promoter causes problems. Eukaryotic RNA Pol II "alone" may not recognize any DNA site at all, and there may well be no universal features of Pol II initiation regions. Nevertheless, we use the term out of habit and convenience. The term promoter clearly implies a function: the region where the site of initiation of transcription is determined (and regulated). However, we must be cautious about using the term to imply a mechanism.

With eukaryotes the term promoter is best understood in the sense of promoter region. And we will often talk about promoter elements: individual features of the promoter region.

TATA

The TATA box is a common Pol II promoter element, and central to most discussion of Pol II. If present, it is located rather close to the start site; it is important for positioning the polymerase properly in front of the start site (Fig 10.28).

> There is some generality that TATA is found in the most active promoters and is absent from some promoters that function at a low level. Weaver expresses this preference another way, p 280.

Inr

Another promoter element is Inr, or initiator element (p 283; also see Fig 10.25 and Fig 11.17). Inr is a site overlapping the transcriptional start site. It's role is not entirely clear, with seemingly different results coming from different systems. Inr and TATA are found in various combinations, and some promoters have neither.

Robinson and Lopes (2000) analyze a very weak yeast promoter, and find an essential Inr element, but no essential TATA box.

Other; enhancers

The rest of the "promoter" -- or "promoter region"? In the broad sense, it is generally considered to include all the sequences that are essential for proper functioning of the gene. Most of the promoter elements are "upstream" of the gene, but some may be downstream (Fig 10-25).

Fig 10-26 to 31 show examples of how promoter elements may be recognized. Also see Sect F, above. That is, we identify sequences necessary for gene expression as promoter elements. (Note how much more complex these experiments are than those which led to early identification of bacterial regulatory elements. In particular note the need for DNA technology.) Individual promoter elements are eventually associated with specific TF or regulatory proteins. Sequence analysis may lead us to associate an element with a known factor (i.e., a factor of known specificity). In vitro biochemical analysis may allow direct identification of the protein that binds to a site. Thus the promoter story merges into the regulatory site story.

> The yeast two hybrid assay is a popular method for identifying protein interactions. It helps you find proteins that interact with a particular protein you have. Weaver introduces it in Fig 14.34. Marcotte et al (1999; Ch 9 FR) explore how to find protein interactions by analyzing the genome.

Enhancers are regulatory elements that have somewhat unusual properties, such as working at "unusually" large distances and regardless of orientation. Nevertheless, it may be best simply to include enhancers among the collection of promoter elements. It is not at all clear that there is any fundamental distinction. (Recall pp 255 ff, and Ch 9 handout Sect D for a brief introduction to enhancers in bacteria, where they are a minor issue.)

Silencers are the opposite of enhancers. As Weaver notes, the difference between a site that stimulates or inhibits transcription may be more a matter of which protein is bound than what or where the site is (p 292).

The important "big story" for now is that the promoter region of eukaryotic genes is complex,

H. Pol II general transcription factors; TBP and TAFs

(Ch 11 Sect 1)

General TFs are proteins that are "always" needed for Pol II transcription. These factors, along with Pol II, are necessary and sufficient for "basal" (unregulated) transcription. They can be considered accessory proteins for Pol II, and are not gene specific.

We now recognize several of these general TFs, understand their roles at least partially, and have a sense that they function in an ordered pathway for assembling an active transcription complex. Fig 11.5 shows a simple version. For a TATA-promoter, TBP along with its

"associated factors" (TAFs) is clearly the central player. It binds to the TATA site; other factors are added one by one, and Pol II is recruited to the complex.

> Caution. The complex of TBP and its TAFs is called TFIID (p 301). It is an important story that the composition of TFIID has been clarified. However, older literature uses the term TFIID rather loosely and variably. TFIID was originally the term for an operationally defined factor, which we now realize is quite complex and variable.

Although the pathway discussed above is important conceptually, there is considerable disagreement on the details of the pathway. The disagreement arises both from increasing knowledge about the details, and probable real differences between systems.

TBP is rather unusual in binding to DNA in the minor groove. It also causes a considerable bend in the DNA (Fig 11.10), which may help to open it up.

Coleman et al (1999) explore one aspect of TBP.

TAFs

TAFs can serve as targets for regulatory proteins. They also can have enzymatic activities (p 309).

Some TAFs are histone analogs. Thus TFIID may look something like a nucleosome (Sect M). Further, the structure of this DNA-binding complex may well be influenced by various regulatory proteins. Thus we can easily imagine how nucleosomes and TFIID may compete for DNA sites, probably under the influence of regulatory proteins.

The role of TAF's varies, and in yeast many genes do not require them at all (p 309 bottom, and following).

Raval et al (2001) show that an activator protein can bypass the requirement for a TAF.

(Some) Other TFs (initiation)

TFIIA stabilizes the initial interaction between TFIID (TBP) and the DNA. It is not always needed. Coleman et al (1999) show that TFIIA increases the availability of TBP. It also interferes with other transcriptional repressors that block TBP.

A key role of TFIIB is to define the position of the start site, relative to TATA.

Andel et al (1999) explore the structure of the TFIID-IIA-IIB complex; also see Sect X. Cf Fig 11.23.

TFIIE activates TFIIH (see below).

Willy et al (2000) think they have a TF that is generally required for transcription from a downstream promoter element.

The Pol II tail; TFIIH

The non-phosphorylated Pol II tail (CTD; recall p 4) is an integral part of the multi-protein initiation complex (above); phosphorylation releases the tail from the complex, thus freeing the polymerase for elongation.

It now seems clear that TFIIH is the factor responsible for phosphorylating the Pol II CTD, thus allowing for "promoter clearance". Fig 11.28.

The TFIIH story is complex. "TFIIH" has been implicated not only in CTD-phosphorylation, but also in DNA repair, including transcription-repair coupling, and in cell cycle control. The factor is now recognized as complex; recent work has just begun to sort out the components of "TFIIH", to find which components are involved in which functions. Le Page et al (2000) and Liu et al (2001) explore the role of TFIIH in the genetic diseases Cockayne syndrome and xeroderma pigmentosum.

Yudkovsky et al (2000) explore what happens to the initiation complex upon the transition to elongation. Some of the proteins move out with the elongation complex, and some remain at the promoter -- ready for reinitiation. Also see Fig 11.33. Recall the σ cycle, Fig 6.16.

I. Variations

We have presented a basic story of how TBP interacting at the TATA box is a key step in the orderly assembly of a Pol II initiation complex. This may be a good frame of reference, but variations are known. Some of those variations are noted here.

TATA-less promoters

Some Pol II promoters lack TATA, as noted above. Interestingly, TBP is still required. One attractive idea is that Inr (p 5 above) and its binding proteins may supply an alternative pathway for recruitment of TBP to the initiation complex. Other proteins and binding sites may also be involved; one way or another, TBP is recruited (Fig 11.17).

As noted in Sect D, TBP is also a player for both Pol I and Pol III. In each case, it is part of a larger complex (SL1, TFIIIB). The two brief subsections on TBP for Pols I and III may be worth reading (pp 329, on SLI, & 334) for perspective.

Thus, TBP appears to be a key player in all eukaryotic transcription, but in different ways. Although the story of TBP (or TFIID) historically focused on its binding to TATA, we now must conclude that the central role of TBP in initiation by all three eukaryotic polymerases does not fundamentally involve its ability to bind TATA. TATA-binding is one aspect of defining the promoter region, but is not itself required.

TBP

It is now clear that there are some alternatives to TBP, known as TRFs = TBP-related factors. This is still a fairly new finding, and work on defining their significance is limited. Perhaps there is a family of TBPs, somewhat analogous to σ factors. Weaver introduces TRF on p 311; see Berk (2000); Zhang et al (2001).

Lee et al (2000) discuss an alternative to TFIID, but still using TBP.

> You may be confused by the hedging on the role of TBP. This reflects the historical development. TBP was discovered as a Pol II TF, needed to bind to TATA. Over time, it was shown that it was also needed when there was no TATA. And it was shown to be involved in Pol I and Pol III transcription. Thus the pendulum swung toward emphasizing the universal requirement for TBP, though perhaps with somewhat different ways of functioning in different cases. Then exceptions were found; so far it would seem these can be accommodated by the idea that we always need TBP or a related factor.

Assembly pathway

It may even be that much of the complex is pre-assembled, and recruited to the TBP as a holoenzyme (p 324).

Perspective

Assembly of a Pol II initiation complex involves many protein-protein and some protein-DNA interactions. It should be obvious that this process is a major target for regulation. Anything that affects these interactions, positively or negatively, will serve to modulate gene function.

J. Elongation factors

(p 322)

As noted above, TFIIH is both an initiation factor and an elongation factor. Le Page et al (2000), mentioned above, actually deal with its role during elongation. TFIIF may also have this dual role.

Weaver introduces TFIIS, which clearly stimulates elongation. It is involved in “pause control” and proofreading, ideas we introduced in the Ch 6 handout. Fig 11.37 shows how backing up and proofreading may occur.

Ferdous et al (2001) have recently shown that a protein complex normally associated with protein degradation is part of the elongating Pol; its role is unclear, but may not involve proteolysis.

K. Specific TF

(Ch 12)

We noted earlier that transcription requires three types of proteins: the Pol itself, general TF and specific TF. Specific TF are those required for a specific gene (or set of genes). They are what we would generally call regulatory proteins, especially activator proteins. This is the major topic of Ch 12, which we will skip.

Lorch et al (2000) update the story of "mediators" (p 371), including their histone acetylating activities.

Bell et al (2001) review insulators (p 369).

Nucifora et al (2001) implicate a TF in Huntington disease, by an unusual mechanism.

L. Effect of chromatin structure -- overview

(Ch 13)

The bacterial chromosome is usually considered to be "free DNA", although that may well be an oversimplification. [Recall Azam et al (1999; Ch 9 FR) for an introduction to proteins that are bound to the bacterial chromosome.] In contrast, eukaryotic chromosomes are known to be complex and highly organized nucleoprotein structures, collectively called chromatin. An important lesson of recent years is that the issue of chromatin vs DNA is a major one for understanding eukaryotic transcription. The simple story is that the normal chromatin structure is rather inhibitory to transcription -- especially to finding genes. But there is more to it than just that.

Weaver discusses chromatin structure and its implications for transcription in Ch 13. The purpose of our coverage here is to briefly introduce you to some key ideas, which you can build on beyond this course.

There are two major topics in Ch 13.

- Sections 1 and 2 discuss the structure of chromatin. We will discuss this just enough to allow discussion of Sect 3.

- Sect 3 deals with the more interesting issue of chromatin function, in particular how the chromatin structure affects its function.

I have broken this down into several sections here. However, there is much interconnection. Some references are noted in the final, miscellaneous section (U).

M. Chromatin structure -- the nucleosome

(Ch 13 Sect 1-2, briefly)

Some key ideas…

There are many levels of chromatin structure. We will focus on structural issues that seem relevant to our discussion of function, below, and will try to minimize discussion of structure per se.

We focus on the nucleosome. A general understanding of the nucleosome at about the level of Fig 13.3c is a good place to start.

Nucleosomal proteins are the highly basic proteins known as histones. Histones of the H2, H3 and H4 classes form the core nucleosome. We will say little about these in detail.

The level immediately above the basic nucleosome is the "linker" -- the H1 histone that somehow "ties down" the nucleosomes. Figs 13.6 and 13.8a give you an idea of where H1 is. Caution: There is still considerable uncertainty about H1; we mention it because of its demonstrated importance.

Higher levels of chromatin structure remain poorly understood, and we will ignore them.

N. Effect of histones on gene activity -- general

(Ch 13 Sect 3, but largely ignore first subsection on 5S RNA transcription, which is from Pol III.)

The key problem that must be solved is that nucleosomes are a barrier to recognizing binding sites on the DNA (for RNAP or for regulatory proteins). (A more neutral statement might be that chromatin and free DNA look different, to proteins.)

Two general types of solutions are recognized…

- In some cases, the nucleosomes are positioned in such a way that they do not block a key recognition site. More about this in Sect O, below.

- The histone proteins are key players in chromatin structure. They are highly basic proteins, which bind tightly to DNA. In general, the histones probably non-specifically tend to repress gene function.

There seem to be two general approaches to overcoming histone-mediated repression:

- modify the histones; Sections P and Q, below.
- nucleosome remodeling machines; Section R, below.

As we discuss these various mechanisms for dealing with chromatin structure, we should emphasize that this field is still in an early stage. We are learning new tricks, and learning how some specific genes are regulated. It is not yet clear what the big generalities are.

O. Nucleosome positioning

(pp 397 ff)

If nucleosomes are positioned so as to avoid the promoter, then clearly nucleosome-mediated repression is avoided.

This raises the issue of how nucleosomes are positioned, since at first view they would seem to be quite non-specific. The best answer to this is probably that some regulatory proteins bind to the DNA before nucleosome assembly, thus "phasing" the nucleosomes and leaving key parts of the DNA "exposed." This implies a competition between regulatory proteins and nucleosomes during formation of new chromatin. The attached Fig, p 28, describes one logic for this happening.

Weaver introduces "nucleosome positioning" in a subsection on pp 397 ff. DNAse hypersensitive sites (p 399) are interpreted as an observable manifestation of nucleosome positioning.

Cirillo & Zaret (1999) describe an example of nucleosome phasing. In this case, the regulatory protein that positions the nucleosomes starts by binding to nucleosomal DNA. That is, nucleosomes are re-positioned by the regulatory protein.

P. Histone modification. 1. acetylation and deacetylation, a simple story

(mostly pp 403-8)

A second approach to reducing histone repression is to modify the histones, so they bind less tightly to DNA. In recent years, we have learned about enzymes that modulate histone-DNA binding by decreasing or increasing the charge at key sites on histones. These proteins acetylate or deacetylate the side chain amino group on lysine; acetylation converts the basic amine to a neutral amide.

The enzymes that acetylate histones are known as histone acetyl transferases (HAT) or (more informally) histone acetylases; the deacetylation is done by histone deacetylases (HDAC).

There are two rather distinct roles for histone acetylation, or the acetylation/deacetylation cycle.

- Newly synthesized histones are acetylated (in the cytoplasm, by type B HATs); they are then deacetylated as part of chromatin assembly.

- In contrast, acetylation/deacetylation of intact chromatin is for gene regulation.

Both chromatin assembly and gene regulation are interesting topics, but the latter is our immediate concern.

Many gene activation events are now understood to be associated with histone acetylation, reducing the charge on the histone tails and thus loosening their grip on the DNA. Deacetylation reverses such activation.

Weaver discusses the roles of acetylation and deacetylation in gene regulation, pp 403-8.

Regions of chromatin with unusually low levels of acetylation are "silenced." Weaver discusses the role of acetylation -- or rather, deacetylation -- in developing silencing of a chromosomal region in the subsection on p 409. Sekinger & Gross (2001) explore how this silencing prevents gene function. How close yeast heterochromatin is to heterochromatin in higher organisms remains an open question.

> The term "silence" is introduced in two somewhat different contexts. In Ch 10 (p 291), silencers are discussed as regulatory sites, probably for specific genes. In Ch 13 (p 409) silencing is discussed as repression of a broad chromosomal region. In both cases, silencing refers to turning genes off, as the term would suggest. In one case the focus is local, in the other case it is less local. It might be best to be open for how close the mechanisms are. It is not yet clear how repressive (or activating) effects spread along the chromosome. Weaver's discussion of "insulators" (Ch 12, p 369) addresses this; we won't discuss this, and it is quite murky.

Q. Histone modification. 2, other; the histone code

[not in book]

Above we suggested that the effect of histone modification (specifically, acetylation) was due to the charge change affecting how tightly the histones bind to DNA. Indeed there is evidence that the charge modification is what is important (Ren & Gorovsky, 2001).

However, another possibility is that the modification affects how other proteins interact with the histones. This is an emerging idea, with increasing support. It even leads to the idea of a "histone code", popularized by Strahl & Allis (2000).

There are other histone modifications, such as methylation, phosphorylation, and ubiquitination. Sorting out what all these mean is an active field. Nakayama et al (2001), Wang et al (2001) and Lo et al (2001) are recent examples showing the interconnected roles of different modifications. Reinke et al (2001) show that the effect of acetylation in their system is due to specificity of interaction with another protein. No one is actually proposing any general simple rules, but it is now clear that one histone modification can lead to another.

Maher (2001) is a good recent introduction. Cheung et al (2000) review the field.

Raval et al (2001) discuss an activator protein that is an acetylase, though it is not clear whether it acts by acetylating histones.

Vogelauer et al (2000) discuss the implications of acetylation for other chromosome processes.

R. Chromatin remodeling

(pp 408-9)

Histone acetylation may loosen the histones, but it still leaves the nucleosome substantially intact. Another layer of dealing with nucleosomal repression is due to "nucleosome remodeling machines", such as the ATP-dependent SWI/SNF or RSC, which somehow disrupt and/or move the nucleosomes.

Weaver's discussion of chromatin remodeling complexes (pp 408-9) is a useful introduction. I would add a few comments:

Although the SWI/SNF chromatin remodeling complex was discovered first (in yeast) and is the most studied, the RSC complex may actually be more important. It is much more abundant, and is essential. It may well be that SWI/SNF and RSC work rather similarly.

It would now seem that the primary function of RSC (and probably also SWI/SNF) is a loosening of the DNA as it enters the nucleosome. Lorch et al (2001).

Weaver makes a distinction between how SWI/SNF and ISWI work. I'm not sure that the distinction he makes really holds up as further information becomes available. In fact, ISWI is probably one part of the SWI/SNF complex. As you study chromatin remodeling complexes more, you will read about SWI/SNF, RSC, and ISWI. Keep an open mind!

> Technically, SWI/SNF refers to a complex from yeast, and ISWI refers to a Drosophila protein. However, both seem to have widespread homologs, and the terms are often used generally.
>
> The terms "nucleosome remodeling machine" and "chromatin remodeling machine" are used interchangeably.

Agalioti et al (2000) discuss the role of remodeling and other events in the ordered formation of an initiation complex for the β-interferon promoter. Bultman et al (2000) is another example of work on chromatin remodeling. The breast cancer genes, BRCA1 and BRCA2 have been implicated in chromosome remodeling, among other things (Scully & Livingston, 2000).

Tyler & Kadonaga (1999) discuss chromatin remodeling in the context of repressive effects.

S. Elongation

(p 410 ff)

The common view is that elongation through nucleosomes is much less of a problem. The work described in the final subsection (p 410), and summarized in Fig 13.46, suggests that nucleosomes are simply moved by an advancing RNAP.

Orphanides & Reinberg (2000) review the issue.

T. Chromosomes and nuclei are dynamic structures

In Ch 1 handout Sect J we introduced a caution… that it is all too easy to think of biological structures as static, perhaps going back and forth between two well-defined states. I guess it is progress that we are now beginning to heed this caution in the nucleus. An emerging view is to realize how dynamic the nucleus is, and how dynamic the chromosome is. In the narrowest context here, it helps to think of histone proteins in equilibrium with the chromosome structure. At any given time, some fraction of the chromosomes may be "more tightly bound", but the situation is dynamic. Regulatory proteins are now thought of as affecting the on/off rates, thus the equilibrium. Thinking of structures as dynamic provides, both literally and figuratively, more flexibility.

Misteli (2001) and a set of articles listed as Wolffe & Hansen (2001) develop these ideas.

U. Chromatin structure -- miscellaneous

The developing story, then, is how chromatin is modified and remodeled to allow transcription. We have learned some of the tools the cell uses to find genes in chromatin, and we have seen some examples of how these interact in an orderly way.

If you find that the collection of facts and examples above seem to lack a unifying theme, you are correct in a sense. One might talk about the broad theme that transcriptional regulation involves modulating the access of RNA Pol to its initiation sites, but beyond that the details vary widely.

Some miscellaneous papers that I have not classified above…

Kornberg & Lorch (1999) is a good introductory and historical overview.

Cirillo & Zaret (1999) discuss a TF that binds specifically to nucleosomal DNA, via phasing but without acetylation.

In a recent review, Struhl (1999) provocatively argues that the logic of regulation is fundamentally different in eukaryotes -- because of the role of chromatin proteins in repressing genes.

A special aspect of chromatin modification is DNA methylation, which is also involved in gene regulation (e.g., Bell et al, 2001).

Roth & Roth (2000) review similar chromosome remodeling and modifications as they seem to be involved in immunoglobulin gene rearrangements.

V. Preview -- Ch 14-16

The next Part of Weaver deals with post-transcriptional events -- modifications of RNA following its primary synthesis. Time constraints (and priority choices) do not allow us to consider this material with any depth in X107A. We will try to leave some time in the last regular class for an overview; you will not be responsible for that on the final exam.

Two post-transcriptional modifications that are common in eukaryotic mRNA will be mentioned and used in the upcoming chapters on protein synthesis. These are the typical modifications of the two ends of the message, to give a 5′-cap structure (Fig 15.4) and a 3′-poly(A) tail (Fig 15.10b). For our discussions of protein synthesis, you do not need to know anything about these, other than that they are distinctive structures.

The recent completion of the human genome has brought new attention to post-transcriptional processing of mRNA. The simple story is that we have about 30,000 genes but make at least three times that many proteins. This implies that we make multiple proteins per gene -- not as a rarity but as a major approach to achieving our protein diversity. The two most likely ways to make multiple proteins from one gene involve alternative splicing and editing, both of which are mRNA processing. Splicing is the subject of Ch 14, and there is a brief subsection introducing alternative splicing, p 448. His contradictory sentences in the first paragraph about its prevalence presumably highlight the impact of the recent gene count for humans. mRNA editing is introduced in Ch 16 Sect 4. However, the one type of editing that Weaver discusses has no known relevance to humans (or to any organisms other than trypanosomes, so far as I know), and he does not talk about the type that is relevant.

W. Further reading

K Struhl, Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. Cell 98:1, 7/9/99. Minireview. A provocative overview of the role of chromatin structure in eukaryotic transcription.

R D Kornberg & Y Lorch, Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell 98:285, 8/6/99. Minireview. Overview of how the nucleosome affects gene function. Some of the review is historical, but they bring the story up to date.

C A Virbasius et al, A human nuclear-localized chaperone that regulates dimerization, DNA binding, and transcriptional activity of bZIP proteins. Mol Cell 4:219, 8/99. One role of chaperone proteins is "holding" regulatory proteins in one or another conformation. Weaver shows an example in Fig 12.7.

R A Coleman et al, TFIIA regulates TBP and TFIID dimers. Mol Cell 4:451, 9/99. A functional study of one role of TFIIA. TBP binding is often rate limiting in initiation -- which is why it is such a good target for gene regulation. It is known that dimerization of TBP reduces transcription. That is, TBP dimers represent a sequestered form of TBP. Here they show that TFIIA interferes with that dimerization. Thus TFIIA increases the availability of TBP. Note that this is only one of the roles of TFIIA.

J K Tyler & J T Kadonaga, The "dark side" of chromatin remodeling: repressive effects on transcription. Cell 99:443, 11/24/99. Minireview. One of a series of minireviews on transcriptional repression in this issue of Cell. The basic points here are to relate chromatin remodeling to the primary process of chromatin assembly, and to show that these processes can cause repression as well as activation.

C Niehrs & N Pollet, Synexpression groups in eukaryotes. Nature 402:483, 12/2/99. A synexpression group -- the authors' term -- is a group of genes that are regulated together. Here they review "group regulation" in eukaryotes. Since eukaryotes lack polycistronic messages and operons, in the traditional sense, such group regulation was recognized slowly.

F Andel et al, Three-dimensional structure of the human TFIID-IIA-IIB complex. Science 286:2153, 12/10/99. (+ related article, Brand, p 2151). Cf Fig 11.23. Also see web link, from Berkeley news; see Sect X.

L A Cirillo & K S Zaret, An early developmental transcription factor complex that is more stable on nucleosome core particles than on free DNA. Mol Cell 4:961, 12/99. They show that a particular TF binds specifically to nucleosomal DNA, and initiates a sequence of events that leads to transcription. Part of the story is nucleosome phasing -- positioning of a nucleosome by a bound protein. This follows on work from Zaret that Weaver discusses, pp 401-27.

B D Strahl & C D Allis, The language of covalent histone modifications. Nature 403:41, 1/6/00. Good review. In addition to acetylation, histones may undergo methylation and phosphorylation. Both of these modifications may be involved in the chromatin remodeling during gene activation. Phosphorylation may also be involved in higher levels of chromosome condensation.

K S Gajiwala et al, Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. Nature 403:916, 2/24/00. This protein binds DNA through two β-strands and a loop ("wing") between them. You can add this DNA-binding protein motif to Weaver's list on p 343. Also recall Ch 9 handout Sect E.

F Le Page et al, Transcription-coupled repair of 8-oxoguanine: requirement for XPG, TFIIH, and CSB and implications for Cockayne syndrome. Cell 101:159, 4/14/00. TFIIH may be the most complex TF. It is involved in initiation and in elongation -- and in DNA repair. Sorting out its many subunits and activities is an ongoing challenge. In terms of TFIIH, this work focuses on its role in DNA repair. TFIIH is involved in transcription-coupled repair (TCR) of oxidative lesions. Failure to repair an 8-oxoguanine in the template DNA blocks transcription by Pol II. Hence defects in TCR, including certain defects in TFIIH genes, lead to transcriptional blockage -- and the genetic disease called Cockayne syndrome.

J F Martínez-García et al, Direct targeting of light signals to a promoter element-bound transcription factor. Science 288:859, 5/5/00. (+ News, Nagatani, p 821.) In plants, some genes are directly activated by light, through the phytochrome photoreceptor proteins. Here they show that the activated phytochrome directly binds to a TF. Thus we have a light-induced conformational change of the transcriptional initiation complex.

K A Robinson & J M Lopes, The promoter of the yeast INO4 regulatory gene: a model of the simplest yeast promoter. J Bacteriol 182:2746-2752, 5/00. Analysis of this simple -- and weak -- promoter shows only two small regions (total 68 bp) that are required. One of these (12 bp) contains the transcriptional start sites, and is absolutely essential. This is apparently an Inr element; there is no TATA box (or at least not an essential one). The 12 bp Inr region is sufficient to provide enough gene function for "normal" growth.

T I Lee et al, Redundant roles for the TFIID and SAGA complexes in global transcription. Nature 405:701, 6/8/00. SAGA was identified as an activator complex with histone acetylase activity. These two complexes share several TAFs. Further, each contains a histone acetylase, and those acetylases seem to be redundant. The idea, then, is that TFIID per se may not be essential, since there are alternatives. (But SAGA still acts through TBP.)

Y Lorch et al, Mediator-nucleosome interaction. Molecular Cell 6(1):197, 7/00. The mediator, as the name suggests, mediates -- the interaction between activators and the Pol. Weaver introduces the mediator at the very end of Ch 12. The big message is that mediators are getting more and more complex. Mammalian mediator complexes may contain 20 proteins, and have a molar mass of one million. Here they identify a subunit of mediator that acetylates histones.

G Orphanides & D Reinberg, RNA polymerase II elongation through chromatin. Nature 407:471, 9/28/00. Review. Recent years have seen much progress in understanding how chromatin is opened up to allow initiation of transcription. How that is extended to allow elongation is less clear. A histone acetylase that is part of elongating RNA Pol II has been characterized; The implication is that this enzyme "eases" the way for RNA Pol during transcription.

A J Berk, TBP-like factors come into focus. Cell 103:5, 9/29/00. Minireview. In Ch 10 a major theme was that TBP is universally required for eukaryotic transcription, by all three classes of RNA Pol. We briefly noted possible exceptions. Now, Berk reviews these -- probably the first review article on such exceptions. The information is still very fragmentary, but roles of TBP-like factors are emerging. Also see Zhang et al (2001).

P Cheung et al, Signaling to chromatin through histone modifications. Cell 103(2):263, 10/13/00. Broad review of all types and roles of histone modifications.

P J Willy et al, A basal transcription factor that activates or represses transcription. Science 290:982, 11/3/00. This factor activates transcription from promoters with a DPE (downstream promoter element), but represses transcription from TATA boxes. Here they show that these are distinct activities of the protein.

N Yudkovsky et al, A transcription reinitiation intermediate that is stabilized by activator. Nature 408:225, 11/9/00. Fig 11.33 shows that some of the initiation complex remains behind at the promoter when Pol II goes into elongation phase. This paper explores that further, and shows that the proteins left behind include some of the gene-specific regulatory proteins. The "left-behind" proteins can promote reinitiation. As so often, a caution that there is not agreement on the details.

T Agalioti et al, Ordered recruitment of chromatin modifying and general transcription factors to the IFN-β promoter. Cell 103:667, 11/10/00. Weaver discusses the sequence of events for one chromatin remodeling event on p 408. This article discusses another. The details are rather different. What is important is that each presents a reasonably coherent view of a series of events leading to gene function.

D B Roth & S Y Roth, Unequal access: Regulating V(D)J recombination through chromatin remodeling. Cell 103(5):699, 11/22/00. Minireview. Chromatin events similar to those we discuss for allowing access to transcription may be involved with other activities. Here, they review work on chromatin remodeling, including histone acetylation, in regulation of the gene arrangements that are part of the development of the immune system.

R Scully & D M Livingston, In search of the tumour-suppressor functions of BRCA1 and BRCA2. Nature 408:429, 11/23/00. Review. The BRCA1 gene was the first gene shown to be associated with some forms of familial breast cancer. It behaves as a tumor suppressor gene, and codes for an extremely large protein, whose precise function has proved elusive. BRCA1 has been implicated in gene regulation, DNA repair and recombination. Both BRCA proteins have been implicated in chromosome remodeling.

M Vogelauer et al, Global histone acetylation and deacetylation in yeast. Nature 408:495, 11/23/00. (+ News, Berger, p 412.) Broader view of acetylation, affecting not only transcription but other chromosome processes such as replication and repair. The news item is particular good for an overview.

J L Riechmann et al, Arabidopsis transcription factors: Genome-wide comparative analysis among eukaryotes. Science 290:2105, 12/15/00. An emerging field is analyzing genomes to

see what TF they encode. Here, they take advantage of the first completed plant genome to compare transcriptional proteins in the major eukaryotic kingdoms (plants, animals, fungi).

S Bultman et al, A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. Molecular Cell 6(6):1287-95, 12/00. Evidence for distinct chromosome remodeling complexes in the mouse.

A C Bell et al, Insulators and boundaries: Versatile regulatory elements in the eukaryotic genome. Science 291:447, 1/19/01. Review. (+ related articles, p 493, p 495.) Insulators (or boundary elements) seem to determine how far a regulatory protein can influence a region of DNA. In some operational sense, they divide the chromosome into functional domains. Bell et al review the work on these rather mysterious but presumably very important regulatory elements. It is possible that chromosome methylation, such as in genomic imprinting, works in part by affecting boundary elements.

Y Lorch et al, RSC unravels the nucleosome. Molecular Cell 7(1):89-95, 1/01. This paper explores the action of RSC. The major conclusion is that it causes loosening of the DNA around the nucleosome, from the ends (of the DNA as it enters the nucleosome) inward. If you think of this DNA as being in some equilibrium, free vs nucleosomal bound, RSC drives the equilibrium toward “free”.

A Raval et al, Transcriptional coactivator, CIITA, is an acetyltransferase that bypasses a promoter requirement for TAF(II)250. Molecular Cell 7(1):105-115, 1/01. CIITA is a tissue-specific activator needed for expression of MHC genes. Basal level of expression of the MHC genes requires the $TAF_{II}250$, but expression with the activator does not require that TAF. The activator is an acetyltransferase; that activity is essential for its activation role. It is not clear what the natural substrate for acetylation is; lab work is done by assaying acetylation of histones. This work illustrates the importance of acetylation, and also is the first example of a cellular activator protein that can bypass a known requirement for $TAF_{II}250$.

T Misteli, Protein dynamics: Implications for nuclear architecture and gene expression. Science 291:843, 2/2/01. Review. A developing view is that nuclear architecture is dynamic, thus shifting the emphasis from describing structures to understanding dynamic interactions. As an example, histone H1 molecules (linker histone) are continuously exchanging in/out of chromatin. Regulatory proteins are better thought of as influencing that equilibrium, rather than as displacing the histones. Technical advances, such as direct microscopy of fluorescence labeled protein in living cells, have played a key role in this shift. For more, see Wolffe & Hansen (2001), below.

J H Liu et al, Defective interplay of activators and repressors with TFIIH in xeroderma pigmentosum. Cell 104(3):353-363, 2/9/01. TFIIH is involved in transcription and in DNA repair. Mutations in the genes for two of its many subunits can lead to cancer. This paper explores the subtleties of interactions of TFIIH with transcriptional regulatory proteins. These proteins can regulate TFIIH-mediated promoter clearance. They suggest that this effect, on specific genes, could account for TFIIH mutations leading to increased risk for cancer.

1) A P Wolffe & J C Hansen, Nuclear visions: Functional flexibility from structural instability. Cell 104(5):631-634, 3/9/01. 2) T Pederson, Protein mobility within the nucleus-

What are the right moves? Cell 104(5):635-638, 3/9/01. 3) S M Gasser, Positions of potential: Nuclear organization and gene expression. Cell 104(5):639-642, 3/9/01. A series of three minireviews. See Misteli (2001), above.

F C Nucifora et al, Interference by Huntingtin and Atrophin-1 with CBP-mediated transcription leading to cellular toxicity. Science 291:2423, 3/23/01. There are several neurodegenerative diseases that involve formation of protein deposits. In each case, one question is how the protein deposit is formed, and another is why it is toxic -- assuming that the deposit is what is toxic. Here they show that two deposit-forming proteins, which aggregate by virtue of extended polyglutamine stretches, co-precipitate a key transcription factor, which contains a short polyglutamine stretch. Overexpression of the TF reduced the neurotoxicity.

H Reinke et al, A transient histone hyperacetylation signal marks nucleosomes for remodeling at the PHO8 promoter in vivo. Molecular Cell 7(3):529-538, 3/01. They provide evidence that the binding of the activator protein to the chromatin targets a histone acetylase to acetylate a specific target nucleosome. This in turn allows the remodeling machinery to loosen the nucleosome. Note in this case that the acetylated histone may act by being a specific target for another protein.

J Nakayama et al, Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science 292:110, 4/6/01. (+ News, Berger, p 64.) They discuss an interaction between two histone modifications, methylation and acetylation. Specifically, they suggest that deacetylation of one lysine leads to methylation of a nearby lysine, which then leads to a stable binding of a silencing protein. Also recall McBride & Silver (2001; Ch 3 handout) for histone methylation.

E A Sekinger & D S Gross, Silenced chromatin is permissive to activator binding and PIC recruitment. Cell 105(3):403-414, 5/4/01. This paper explores how the silencing system Weaver describes in Fig 13.39 works. Surprisingly, they show that the silenced DNA still binds activator proteins, TBP and Pol II. I wonder whether this result will hold up.

D Zhang et al, Spermiogenesis deficiency in mice lacking the Trf2 gene. Science 292:1153, 5/11/01. TRF = TBP-related factors, i.e., alternatives to TBP. This is some of the first work showing a role for a mammalian TRF. Recall Berk (2000).

W Hausner & M Thomm, Events during initiation of Archaeal transcription: open complex formation and DNA-protein interactions. J Bacteriol 183:3025-3031, 5/01. The molecular biology of the Archaea has features of both prokaryotic and eukaryotic systems. For transcription, there are a TATA box and a TBP, and a TFIIB, but not other typical eukaryotic factors. Here they show that formation of the open complex does not require ATP hydrolysis, in contrast to typical eukaryotic transcription.

A Ferdous et al, The 19S regulatory particle of the proteasome is required for efficient transcription elongation by RNA polymerase II. Molecular Cell 7(5):981-991, 5/01. The proteasome is a large complex for protein degradation. Perhaps surprisingly, part of it seems to be associated with elongating RNA Pol II, and mutations that affect these proteins interfere with transcriptional elongation. However, inhibition of the protease activity does not affect

transcription, suggesting that the role of proteasomal proteins in Pol is not as an active protease.

O Calvo & J L Manley, Evolutionarily conserved interaction between CstF-64 and PC4 links transcription, polyadenylation, and termination. Molecular Cell, 7(5):1013-1023, 5/01. That linking occurs on the CTD.

A Klug, Structural biology: A marvellous machine for making messages. Science 292:1844, 6/8/01. News, to accompany two major articles in this issue on structural analysis of Pol II.

Q Ren & M A Gorovsky, Histone H2A.Z acetylation modulates an essential charge patch. Molecular Cell 7(6):1329-1335, 6/01. We often say that the relevance of acetylation of histones is that it affects the charge. Here they show that, for a variant histone in Tetrahymena. (The alternative is that the acetylation was required for some specific recognition. That may be true for some modifications, but not here; Sect Q)

C Wang et al, TAK1 is a ubiquitin-dependent kinase of MKK and IKK. Nature 412:346, 7/19/01. (+ News, Finley, p 283.) Ubiquitin is a protein used to tag proteins in the cell. The ubiquitin tag usually means “degrade me”, but examples of its use in gene regulation are being uncovered. In some cases, ubiquitin tags an inhibitor for degradation, thus leading to gene activation but within its normal role. Here they show an example of how the ubiquitin tag marks a protein for another modification, but without any degradation.

W-S Lo et al, Snf1--a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription. Science 293:1142, 8/10/01. An example of how multiple histone modifications may be coordinated. In this case, phosphorylation obligatorily precedes acetylation during gene activation.

B A Maher, Researchers focus on histone code. The Scientist, Sept 17, 2001, p 15. A brief and readable overview of the emerging idea that histone modifications change the face that histones present to other proteins. (The Scientist is available online.)

X. Computer resources

See my web page for Molecular Biology Internet Resources.

Y. Homework

Weaver Ch 10

#1-5 deal with the three types of eukaryotic RNA Pol, and #22-26 deal focus on Pol I and III; thus these two groups of questions are low priority.

Among the “better” questions… #16-20; 27-28. The answer section below contains some discussion of all of those.

⇒ We will discuss the following two from Weaver, with my additions/modifications (we can discuss others if you wish):

For #19, add part b:

b. One of the two common effects of deleting TATA (main part of question) is quite variable. Which one, and why? In discussing why, the issue is why the TATA box seems critical in some cases and rather nonessential in others. In general terms, how can this be?

For #27, you should propose two experiments, which test different features of what an enhancer is about.

My questions (Most are these are "general perspective" questions. Some go beyond class material, and call on your general knowledge.)

1. Which of the three eukaryotic RNA polymerases transcribes the largest number of genes? Which makes the largest amount of RNA? Why was so much early work on eukaryotic transcription done with rRNA?

2. Is there anything fundamental about the use of three distinct RNA polymerases in eukaryotes? In other words, is there any particular reason why having three polymerases is a distinct advantage?

3. As we talk about the formation of the initiation complex for Pol II transcription, the big picture would seem to be a multitude of protein-protein interactions along with some protein-DNA interactions. Recently we have recognized that there are enzymatic activities in the initiation complex (in addition to polymerization per se). List as many of these enzyme activities as possible, and describe the role of each.

4. Strubin & Struhl (Cell 68:721, 2/21/92) isolated mutants of TBP that bound well to a mutant TATA site. The mutants carried two essential mutations within a 12 amino acid sequence. Assume (as was true) that they had a good direct selection for TBP function at the mutant site. Why would it be unlikely to get this type of mutant TBP by ordinary mutagenesis, even with a very high degree of mutagenesis? Suggest a solution.

5. As we try to get some conceptual handles on the very complex eukaryotic genome, it is tempting to suggest that, in a given cell at a given time, there are three kinds of genes: "permanently" inactivated, activatable, and active. What can you say about these three classes? What are the relative amounts of these three kinds of genes in Escherichia coli, Saccharomyces cerevisiae, and Homo sapiens? Is it important in discussing this that E. coli is a prokaryote?

6. One interesting part of the TFIID story surfaced when it was found that the cloned TFIID did not behave the same as the "natural" factor. We now understand this particular discrepancy in terms of the TBP story. (That is, what got cloned was TBP, not TFIID.) In

general, what are possible reasons for such a discrepancy when a protein is cloned? That is, why might a cloned protein behave differently than expected from the biochemical work that preceded the cloning?

7. How many proteins are needed to initiate one simple transcriptional event -- in prokaryotes? in eukaryotes? List them. (By “proteins”, the intent is to count distinct protein chains. If a protein contains multiple kinds of subunits, count each one.)

8. What are the important differences in transcription and transcriptional regulation between prokaryotes and eukaryotes?

Z. Partial answers

Weaver Ch 10

16. Remember that there is no "typical" Pol II promoter; they vary widely. But one way or another, they must bind Pol, and that binding is modulated by regulatory proteins. The regulatory proteins "always" include activators.

17. The best generality is that TATA boxes are usually not found in constitutive "housekeeping" genes. Weaver says more on p 280, but not sure how useful it is.

18. The key idea is that the TATA box determines the start site, at lest approximately. Thus if 10 bases are removed, the start will now occur at whatever base is now about 30 bases from the TATA; the precise position is Pol preference, but the general distance is due to spacing from TATA.

19. Altered start site; reduced transcription level. Both of these reflect the basic roles of TATA that we understand.

b (my added part). As Weaver notes with a variety of experimental data, deleting TATA can cause anything from a 0% to 100% reduction in transcription level. TATA (via TBP) plays a central role in forming the initiation complex. However, there are alternative ways to do so. Thus the importance of TATA per se may range from major to minor, depending on the extent to which others players can compensate. We must conclude that the TATA box does not play any universally critical unique role.

At one level, we now understand that other proteins can position TBP. At this level, we accept that TATA box is non-essential, but TBP is essential. More recently, the essentiality of TBP is being called into question, but this is a very immature story at this time.

20. See Fig 10.30. The key idea is to change specific parts of the sequence being tested, but without changing the overall length (= spacing of elements).

27. Enhancers are sites for positive control protein, and are substantially position-independent. To test… Remove the site, and see if gene expression is reduced. Move the site "substantially" or invert it; there should be little effect.

28. The regulatory proteins that act at such sites may be (and often are) tissue-specific.

My questions

1. Clearly, RNA polymerase II transcribes the largest number of genes. However, the few genes transcribed by the other polymerases are transcribed at extremely high levels. Pol I,

which makes rRNA, makes the largest amount of RNA. Early work was done with rRNA for convenience (because there is so much of it), before we knew it mattered.

2. Use of separate polymerases allows specialization (as discussed in #1). Advantages are speculative. With one polymerase, at worst you would need some distinct accessory proteins, e.g., TF or σ factors.

3.

- Kinase activity in a TAF. (p 309)
- Histone acetylase activity in a TAF; presumably to loosen up chromatin. (p 309; p 406)
- Other histone-modifying activities (Sect Q, above).
- Kinase activity of TFIIH, presumably to release CTD of Pol II and allow it to move away from promoter. (Figs 11.28 & 29)
- Helicase activity of TFIIH; presumably to help melt DNA, but is not needed under some conditions. (p 319)
- ATP-dependent chromatin remodeling activity of proteins such as the SWI/SNF complex. (Sect L, above).

I am not concerned that you remember the details of this list. The purpose of the question is to make the general point that the transcription apparatus involves not only "binding" (weak bond) reactions but also covalent bond reactions. (You should know generally about the types of enzymatic reactions that are involved, but remembering which factor provides them is not important.)

If you notice any additions/corrections to this list, from Weaver or my handout, please let me know.

4. Ordinary mutagenesis is random. It would be rare to get two desired mutations within a very small region without getting a lot of extraneous mutations. The "desired" mutation frequency in this case is 1 per 6 aa!

Solutions? The logical point is to <u>focus</u> the mutagenesis, so that only the <u>target</u> region (which was known, or at least suspected) gets mutated. Random mutagenesis of a small restriction fragment, which is subsequently returned to an otherwise untouched gene, is one possibility. Another alternative is chemical synthesis of the target region, using a mixture of bases at specific positions. This is now a common method, and is what they did here.

5. At one level, this question can be answered descriptively. For example, heterochromatin is highly condensed DNA which is unavailable for transcription or for activation. The idea that TF may be part of the chromatin provides a new level of sophistication. We now imagine that committed pre-initiation complexes are a distinct chromatin state, of activatable DNA. (Chromatin that carries such activating factors is either active or at least potentially active, depending on whether additional regulatory factors are required and/or present.)

The three organisms have increasing genome size and increasing developmental complexity (in the order listed). Developmental genes are only used occasionally [in contrast to general metabolism ("housekeeping") genes]. It's clear that only a very tiny fraction of the human genome is active in a given type of cell, thus most is inactivated. This trend goes with genome

size. If you compare bacteria with different degrees of developmental complexity, again there is a trend of genome size with complexity. It's not clear that there is a prokaryotic analog of "inactivated" chromatin.

6. Unfortunately, this is an important question. There are many possibilities… The cloned gene may be the wrong gene. This may be due to a gross error, such as using an incorrect protein sequence. On the other hand, it may be the wrong gene, but one that is related; this may even turn out to be interesting. Another class of possibilities is that the cloning host produces a different form of the protein. This might be due to incorrect translation (probably rare) or to different processing. The lack of glycosylation in bacteria is a common concern when cloning mammalian proteins. The TFIID example illustrates a third class of explanations, that the two preparations differ in purity. In this case, the native protein -- even when "highly purified" -- contains other tightly bound proteins, which are relevant to its function.

7. For prokaryotes, it would seem that all you need is one complete RNA Pol, with a σ. That would be 4 proteins, or 5 chains: $\alpha_2\beta\beta'\sigma$. In other words, it takes 4 genes to code for the basic transcriptional initiation apparatus.

For eukaryotes, the answer is "many".

8. This is intended as a wrap-up question, a chance for you to review and organize a large amount of information. I have no particular answer in mind. We may discuss it in class -- but it's really a personal matter; spend some time thinking about it.

Figure:

From: M Grunstein, Histones as regulators of Genes. Scientific American, October 1992, p 68.

This Fig is somewhat similar in intent to text Figs 13.18, 23, 24.

x107a\wv10_13h
11/6/06