NUCLEOSOMES

5.20.1 Introduction

Key Terms

The **nucleosome** is the basic structural subunit of chromatin, consisting of ~200 bp of DNA and an octamer of histone proteins.

Histones are conserved DNA-binding proteins that form the basic subunit of chromatin in eukaryotes. Histones H2A, H2B, H3, H4 form an octameric core around which DNA coils to form a nucleosome. Histone H1 is external to the nucleosome.

A **nonhistone** is any structural protein found in a chromosome except one of the histones.

Chromatin has a compact organization in which most DNA sequences are structurally inaccessible and functionally inactive. Within this mass is the minority of active sequences. What is the general structure of chromatin, and what is the difference between active and inactive sequences? The high overall packing ratio of the genetic material immediately suggests that DNA cannot be directly packaged into the final structure of chromatin. There must be *hierarchies* of organization.

The fundamental subunit of chromatin has the same type of design in all eukaryotes. The **nucleosome** contains ~200 bp of DNA, organized by an octamer of small, basic proteins into a bead-like structure. The protein components are **histones**. They form an interior core; the DNA lies on the surface of the particle. Nucleosomes are an invariant component of euchromatin and heterochromatin in the interphase nucleus, and of mitotic chromosomes. The nucleosome provides the first level of organization, giving a packing ratio of ~6. Its components and structure are well characterized.

The second level of organization is the coiling of the series of nucleosomes into a helical array to constitute the fiber of diameter ~30 nm that is found in both interphase chromatin and mitotic chromosomes (see **Figure 19.11**). In chromatin this brings the packing ratio of DNA to ~40. The structure of this fiber requires additional proteins, but is not well defined.

The final packing ratio is determined by the third level of organization, the packaging of the 30 nm fiber itself. This gives an overall packing ratio of ~1000 in euchromatin, cyclically interchangeable with packing into mitotic chromosomes to achieve an overall ratio of ~10,000. Heterochromatin generally has a packing ratio ~10,000 in both interphase and mitosis.

We need to work through these levels of organization to characterize the events involved in cyclical packaging, replication, and transcription. We assume that association with additional proteins, or modifications of existing chromosomal proteins, are involved in changing the structure of chromatin. We do not know the individual targets for controlling cyclical packaging. Both replication and

transcription require unwinding of DNA, and thus must involve an unfolding of the structure that allows the relevant enzymes to manipulate the DNA. This is likely to involve changes in all levels of organization.

When chromatin is replicated, the nucleosomes must be reproduced on both daughter duplex molecules. As well as asking how the nucleosome itself is assembled, we must inquire what happens to other proteins present in chromatin. Since replication disrupts the structure of chromatin, it both poses a problem for maintaining regions with specific structure and offers an opportunity to change the structure.

The mass of chromatin contains up to twice as much protein as DNA. Approximately half of the protein mass is accounted for by the nucleosomes. The mass of RNA is < 10% of the mass of DNA. Much of the RNA consists of nascent transcripts still associated with the template DNA.

The **nonhistones** include all the proteins of chromatin except the histones. They are more variable between tissues and species, and they comprise a smaller proportion of the mass than the histones. They also comprise a much larger number of proteins, so that any individual protein is present in amounts much smaller than any histone.

The functions of nonhistone proteins include control of gene expression and higher-order structure. So RNA polymerase may be considered to be a prominent nonhistone. The HMG (high-mobility group) proteins comprise a discrete and well-defined subclass of nonhistones (at least some of which are transcription factors). A major problem in working with other nonhistones is that they tend to be contaminated with other nuclear proteins, and so far it has proved difficult to obtain those nonhistone proteins responsible for higher-order structures.

NUCLEOSOMES

5.20.2 The nucleosome is the subunit of all chromatin

Key Terms

Micrococcal nuclease is an endonuclease that cleaves DNA; in chromatin, DNA is cleaved preferentially between nucleosomes.

A **core histone** is one of the four types (H2A, H2B, H3, H4) found in the core particle derived from the nucleosome (this excludes histone H1).

Key Concepts

- Micrococcal nuclease releases individual nucleosomes from chromatin as 11S particles.
- A nucleosome contains ~200 bp of DNA, 2 copies of each core histone (H2A, H2B, H3, H4), and 1 copy of H1.
- DNA is wrapped around the outside surface of the protein octamer.

When interphase nuclei are suspended in a solution of low ionic strength, they swell and rupture to release fibers of chromatin. **Figure 20.1** shows a lysed nucleus in which fibers are streaming out. In some regions, the fibers consist of tightly packed material, but in regions that have become stretched, they can be seen to consist of discrete particles. These are the nucleosomes. In especially extended regions, individual nucleosomes are connected by a fine thread, a free duplex of DNA. *A continuous duplex thread of DNA runs through the series of particles.*

Figure 20.1 Chromatin spilling out of lysed nuclei consists of a compactly organized series of particles. The bar is 100 nm. Photograph kindly provided by Pierre Chambon.

Individual nucleosomes can be obtained by treating chromatin with the endonuclease **micrococcal nuclease**. It cuts the DNA thread at the junction between nucleosomes. First, it releases groups of particles; finally, it releases single nucleosomes. Individual nucleosomes can be seen in **Figure 20.2** as compact particles. They sediment at ~11S.

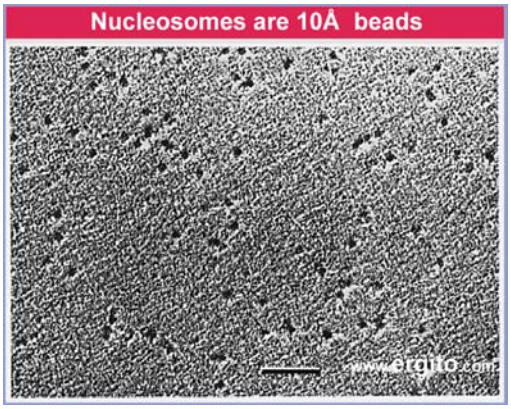

Figure 20.2 Individual nucleosomes are released by digestion of chromatin with micrococcal nuclease. The bar is 100 nm. Photograph kindly provided by Pierre Chambon.

The nucleosome contains ~200 bp of DNA associated with a histone octamer that

consists of two copies each of H2A, H2B, H3, and H4. These are known as the **core histones**. Their association is illustrated diagrammatically in **Figure 20.3**. This model explains the stoichiometry of the core histones in chromatin: H2A, H2B, H3, and H4 are present in equimolar amounts, with 2 molecules of each per ~200 bp of DNA (598; for review see 200).

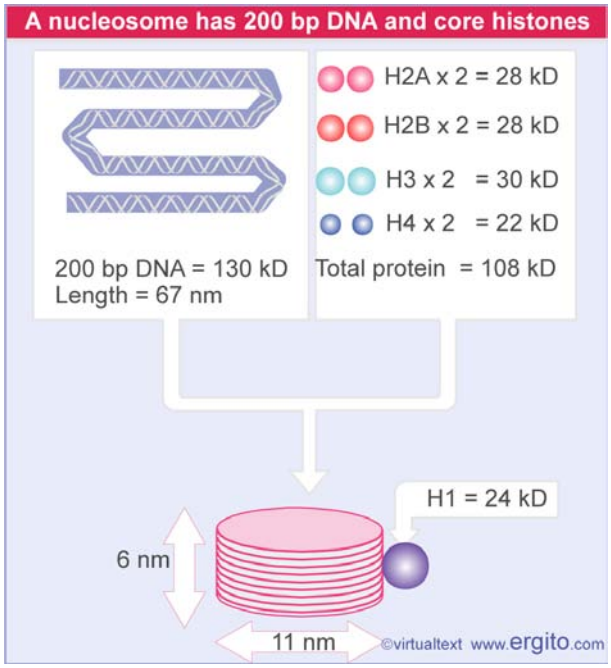

Figure 20.3 The nucleosome consists of approximately equal masses of DNA and histones (including H1). The predicted mass of the nucleosome is 262 kD.

Histones H3 and H4 are among the most conserved proteins known. This suggests that their functions are identical in all eukaryotes. The types of H2A and H2B can be recognized in all eukaryotes, but show appreciable species-specific variation in sequence.

Histone H1 comprises a set of closely related proteins that show appreciable variation between tissues and between species (and are absent from yeast). The role of H1 is different from the core histones. It is present in half the amount of a core histone and can be extracted more readily from chromatin (typically with dilute salt [0.5 M] solution). *The H1 can be removed without affecting the structure of the nucleosome, which suggests that its location is external to the particle.*

The shape of the nucleosome corresponds to a flat disk or cylinder, of diameter 11 nm and height 6 nm. The length of the DNA is roughly twice the ~34 nm circumference of the particle. The DNA follows a symmetrical path around the octamer. **Figure 20.4** shows the DNA path diagrammatically as a helical coil that makes two turns around the cylindrical octamer. Note that the DNA "enters" and "leaves" the nucleosome at points close to one another. Histone H1 may be located in this region (see *Molecular Biology 5.20.4 Nucleosomes have a common structure*) (600).

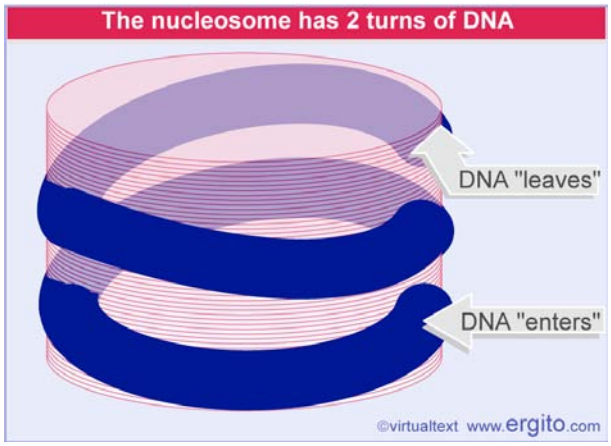

Figure 20.4 The nucleosome may be a cylinder with DNA organized into two turns around the surface.

Considering this model in terms of a cross-section through the nucleosome, in **Figure 20.5** we see that the two circumferences made by the DNA lie close to one another. The height of the cylinder is 6 nm, of which 4 nm is occupied by the two turns of DNA (each of diameter 2 nm).

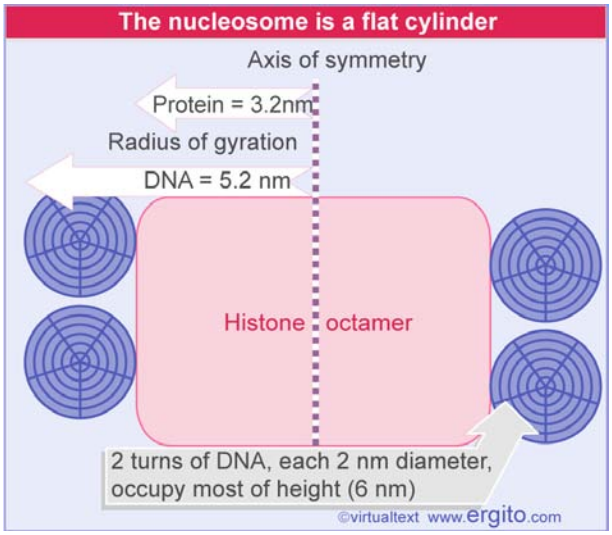

Figure 20.5 The two turns of DNA on the nucleosome lie close together.

The pattern of the two turns has a possible functional consequence. Since one turn around the nucleosome takes ~80 bp of DNA, two points separated by 80 bp in the free double helix may actually be close on the nucleosome surface, as illustrated in **Figure 20.6** (for review see 201).

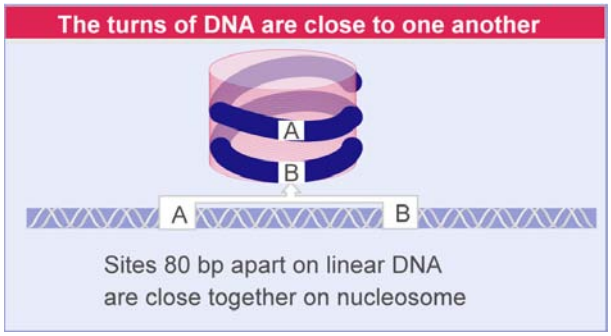

Figure 20.6 Sequences on the DNA that lie on different turns around the nucleosome may be close together.

Reviews

200. Kornberg, R. D. (1977). *Structure of chromatin*. Annu. Rev. Biochem. 46, 931-954.

201. McGhee, J. D., and Felsenfeld, G. (1980). *Nucleosome structure*. Annu. Rev. Biochem. 49, 1115-1156.

References

598. Kornberg, R. D. (1974). *Chromatin structure: a repeating unit of histones and DNA*. Science 184, 868-871.

600. Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D., and Klug, A. (1984). *Structure of the nucleosome core particle at 7* Å *resolution*. Nature 311, 532-537.

NUCLEOSOMES

5.20.3 DNA is coiled in arrays of nucleosomes

Key Concepts

- >95% of the DNA is recovered in nucleosomes or multimers when micrococcal nuclease cleaves DNA of chromatin.
- The length of DNA per nucleosome varies for individual tissues in a range from 154-260 bp.

When chromatin is digested with the enzyme micrococcal nuclease, the DNA is cleaved into integral multiples of a unit length. Fractionation by gel electrophoresis reveals the "ladder" presented in **Figure 20.7**. Such ladders extend for ~10 steps, and the unit length, determined by the increments between successive steps, is ~200 bp.

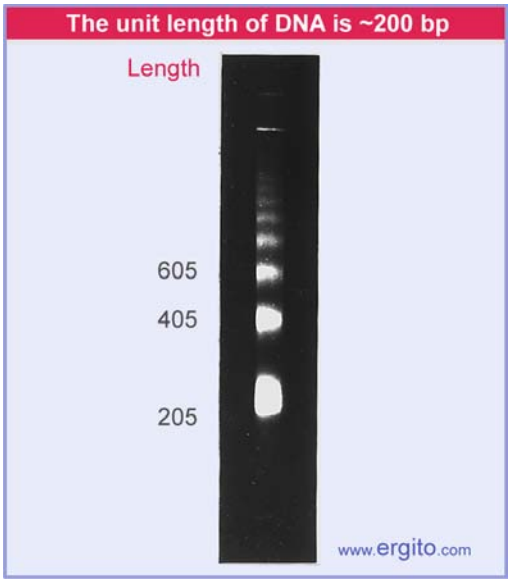

Figure 20.7 Micrococcal nuclease digests chromatin in nuclei into a multimeric series of DNA bands that can be separated by gel electrophoresis. Photograph kindly provided by Markus Noll.

Figure 20.8 shows that the ladder is generated by groups of nucleosomes. When nucleosomes are fractionated on a sucrose gradient, they give a series of discrete peaks that correspond to monomers, dimers, trimers, etc. When the DNA is extracted from the individual fractions and electrophoresed, each fraction yields a band of DNA whose size corresponds with a step on the micrococcal nuclease ladder. The monomeric nucleosome contains DNA of the unit length, the nucleosome dimer contains DNA of twice the unit length, and so on (599).

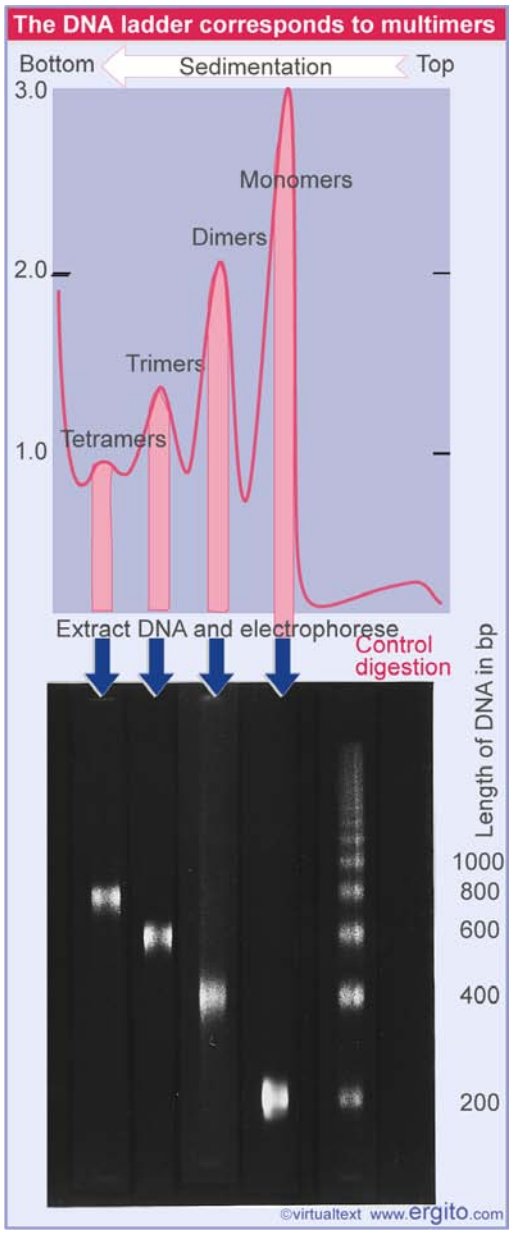

Figure 20.8 Each multimer of nucleosomes contains the appropriate number of unit lengths of DNA. Photograph kindly provided by John Finch.

So each step on the ladder represents the DNA derived from a discrete number of nucleosomes. *We therefore take the existence of the 200 bp ladder in any chromatin to indicate that the DNA is organized into nucleosomes.* The micrococcal ladder is generated when only ~2% of the DNA in the nucleus is rendered acid-soluble (degraded to small fragments) by the enzyme. *So a small proportion of the DNA is specifically attacked; it must represent especially susceptible regions.*

When chromatin is spilled out of nuclei, we often see a series of nucleosomes connected by a thread of free DNA (the beads on a string). However, the need for tight packaging of DNA *in vivo* suggests that probably there is usually little (if any)

free DNA.

This view is confirmed by the fact that *>95% of the DNA of chromatin can be recovered in the form of the 200 bp ladder*. Almost all DNA must therefore be organized in nucleosomes. In their natural state, nucleosomes are likely to be closely packed, with DNA passing directly from one to the next. Free DNA is probably generated by the loss of some histone octamers during isolation.

The length of DNA present in the nucleosome varies somewhat from the "typical" value of 200 bp. The chromatin of any particular cell type has a characteristic average value (±5 bp). The average most often is between 180 and 200, but there are extremes as low as 154 bp (in a fungus) or as high as 260 bp (in a sea urchin sperm). The average value may be different in individual tissues of the adult organism. And there can be differences between different parts of the genome in a single cell type. Variations from the genome average include tandemly repeated sequences, such as clusters of 5S RNA genes.

References

599. Finch, J. T. et al. (1977). *Structure of nucleosome core particles of chromatin*. Nature 269, 29-36.

NUCLEOSOMES

5.20.4 Nucleosomes have a common structure

Key Concepts

- Nucleosomal DNA is divided into the core DNA and linker DNA depending on its susceptibility to micrococcal nuclease.
- The core DNA is the length of 146 bp that is found on the core particles produced by prolonged digestion with micrococcal nuclease.
- Linker DNA is the region of 8-114 bp that is susceptible to early cleavage by the enzyme.
- Changes in the length of linker DNA account for the variation in total length of nucleosomal DNA.
- H1 is associated with linker DNA and may lie at the point where DNA enters and leaves the nucleosome.

A common structure underlies the varying amount of DNA that is contained in nucleosomes of different sources. The association of DNA with the histone octamer forms a core particle containing 146 bp of DNA, irrespective of the total length of DNA in the nucleosome. The variation in total length of DNA per nucleosome is superimposed on this basic core structure.

The core particle is defined by the effects of micrococcal nuclease on the nucleosome monomer. The initial reaction of the enzyme is to cut between nucleosomes, but if it is allowed to continue after monomers have been generated, then it proceeds to digest some of the DNA of the individual nucleosome. This occurs by a reaction in which DNA is "trimmed" from the ends of the nucleosome.

The length of the DNA is reduced in discrete steps, as shown in **Figure 20.9**. With rat liver nuclei, the nucleosome monomers initially have 205 bp of DNA. Then some monomers are found in which the length of DNA has been reduced to ~165 bp. Finally this is reduced to the length of the DNA of the core particle, 146 bp. (The core is reasonably stable, but continued digestion generates a "limit digest", in which the longest fragments are the 146 bp DNA of the core, while the shortest are as small as 20 bp.)

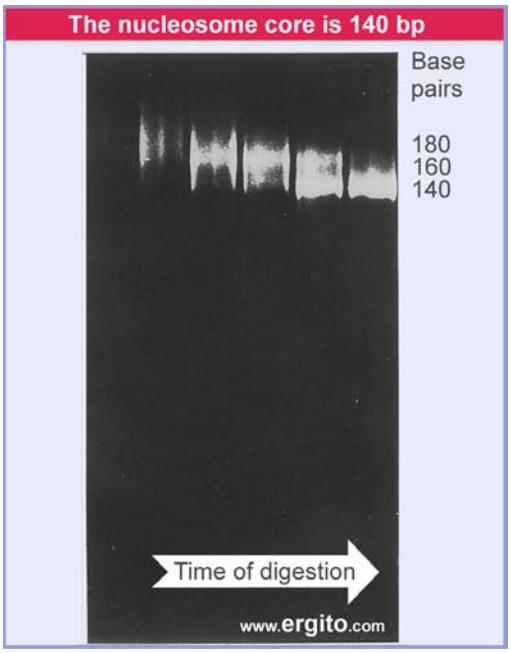

Figure 20.9 Micrococcal nuclease reduces the length of nucleosome monomers in discrete steps. Photograph kindly provided by Roger Kornberg.

This analysis suggests that the nucleosomal DNA can be divided into two regions:

- Core DNA has an invariant length of 146 bp, and is relatively resistant to digestion by nucleases.
- Linker DNA comprises the rest of the repeating unit. Its length varies from as little as 8 bp to as much as 114 bp per nucleosome.

The sharp size of the band of DNA generated by the initial cleavage with micrococcal nuclease suggests that the region immediately available to the enzyme is restricted. It represents only part of each linker. (If the entire linker DNA were susceptible, the band would range from 146 bp to >200 bp.) But once a cut has been made in the linker DNA, the rest of this region becomes susceptible, and it can be removed relatively rapidly by further enzyme action. The connection between nucleosomes is represented in **Figure 20.10**.

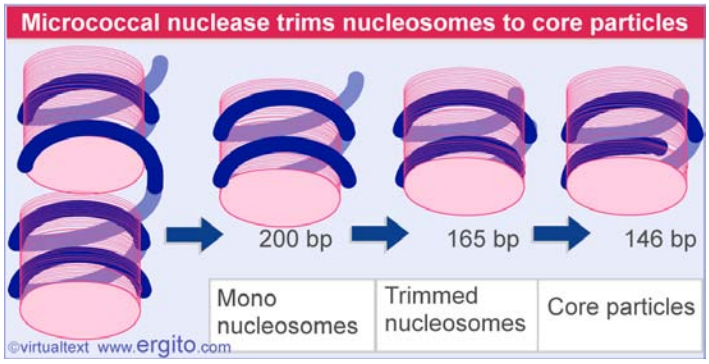

Figure 20.10 Micrococcal nuclease initially cleaves between nucleosomes. Mononucleosomes typically have ~200 bp DNA. End-trimming reduces the length of DNA first to ~165 bp, and then generates core particles with 146 bp.

Core particles have properties similar to those of the nucleosomes themselves, although they are smaller. Their shape and size are similar to nucleosomes, which suggests that the essential geometry of the particle is established by the interactions between DNA and the protein octamer in the core particle. Because core particles are more readily obtained as a homogeneous population, they are often used for structural studies in preference to nucleosome preparations. (Nucleosomes tend to vary because it is difficult to obtain a preparation in which there has been no end-trimming of the DNA.)

What is the physical nature of the core and the linker regions? *These terms were introduced as operational definitions that describe the regions in terms of their relative susceptibility to nuclease treatment.*This description does not make any implication about their actual structure. However, it turns out that the major part of the core DNA is tightly curved on the nucleosome, whereas the terminal regions of the core and the linker regions are more extended (see *Molecular Biology 5.20.5 DNA structure varies on the nucleosomal surface*).

The existence of linker DNA depends on factors extraneous to the four core histones. Reconstitution experiments *in vitro*show that histones have an intrinsic ability to organize DNA into core particles, but do not form nucleosomes with the proper unit length. The degree of supercoiling of the DNA is an important factor. Histone H1 and/or nonhistone proteins influence the length of linker DNA associated with the histone octamer in a natural series of nucleosomes. And "assembly proteins" that are not part of the nucleosome structure are involved *in vivo* in constructing nucleosomes from histones and DNA (see *Molecular Biology 5.20.10 Reproduction of chromatin requires assembly of nucleosomes*).

Where is histone H1 located? The H1 is lost during the degradation of nucleosome monomers. It can be retained on monomers that still have 165 bp of DNA; but is always lost with the final reduction to the 146 bp core particle. This suggests that H1 could be located in the region of the linker DNA immediately adjacent to the core DNA.

If H1 is located at the linker, it could "seal" the DNA in the nucleosome by binding at the point where the nucleic acid enters and leaves (see **Figure 20.4**). The idea that

H1 lies in the region joining adjacent nucleosomes is consistent with old results that H1 is removed the most readily from chromatin, and that H1-depleted chromatin is more readily "solubilized". And it is easier to obtain a stretched-out fiber of beads on a string when the H1 has been removed (618).

Last updated on 7-21-2003

References

618. Shen, X. et al. (1995). *Linker histones are not essential and affect chromatin condensation in vitro*. Cell 82, 47-56.

NUCLEOSOMES

5.20.5 DNA structure varies on the nucleosomal surface

Key Concepts

- 1.65 turns of DNA are wound round the histone octamer.
- The structure of the DNA is altered so that it has an increased number of base pairs/turn in the middle, but a decreased number at the ends.

The exposure of DNA on the surface of the nucleosome explains why it is accessible to cleavage by certain nucleases. The reaction with nucleases that attack single strands has been especially informative. The enzymes DNAase I and DNAase II make single-strand nicks in DNA; they cleave a bond in one strand, but the other strand remains intact at this point. So no effect is visible in the double-stranded DNA. But upon denaturation, short fragments are released instead of full-length single strands. If the DNA has been labeled at its ends, the end fragments can be identified by autoradiography as summarized in **Figure 20.11**. (This is exactly analogous to the restriction mapping technique shown in **Figure S 37**.)

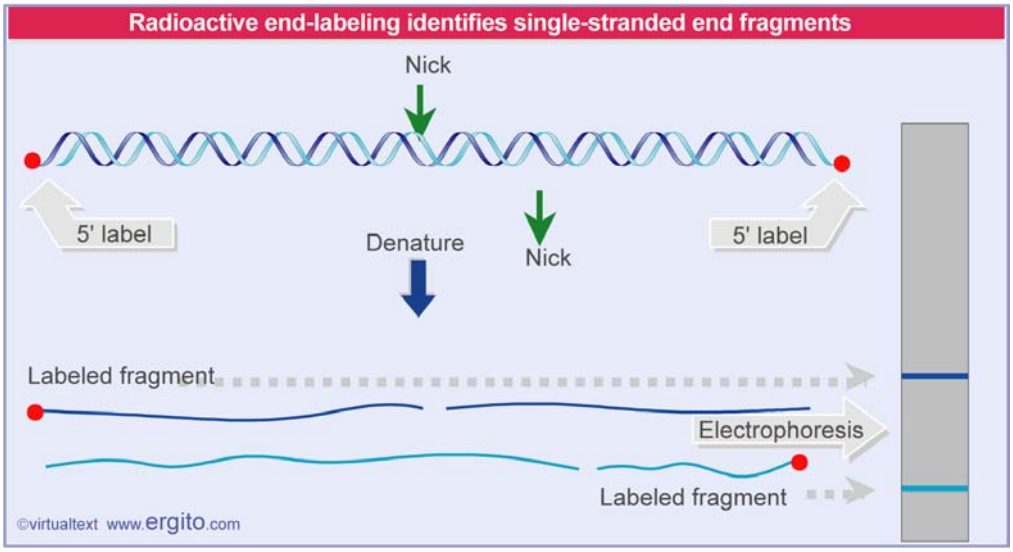

Figure 20.11 Nicks in double-stranded DNA are revealed by fragments when the DNA is denatured to give single strands. If the DNA is labeled at (say) 5 ′ ends, only the 5 ′ fragments are visible by autoradiography. The size of the fragment identifies the distance of the nick from the labeled end.

When DNA is free in solution, it is nicked (relatively) at random. The DNA on nucleosomes also can be nicked by the enzymes, *but only at regular intervals.*When the points of cutting are determined by using radioactively end-labeled DNA and then DNA is denatured and electrophoresed, a ladder of the sort displayed in **Figure 20.12** is obtained.

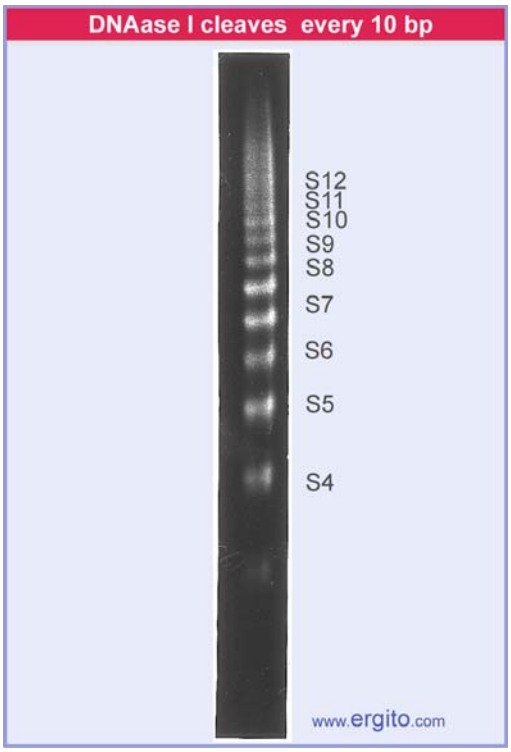

Figure 20.12 Sites for nicking lie at regular intervals along core DNA, as seen in a DNAase I digest of nuclei. Photograph kindly provided by Leonard Lutter.

The interval between successive steps on the ladder is 10-11 bases. The ladder extends for the full distance of core DNA. The cleavage sites are numbered as S1 through S13 (where S1 is ~10 bases from the labeled 5′ end, S2 is ~20 bases from it, and so on). Their positions relative to the DNA superhelix are illustrated in **Figure 20.13**.

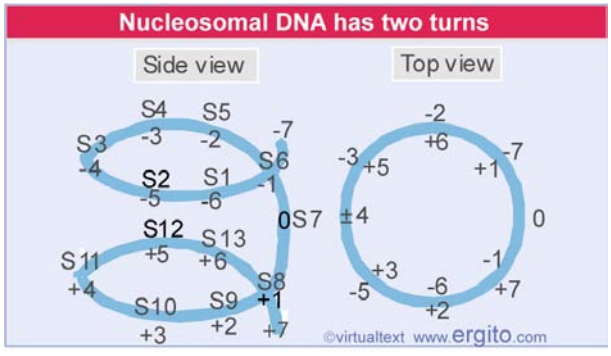

Figure 20.13 Two numbering schemes divide core particle DNA into 10 bp segments. Sites may be numbered S1 to S13 from one end; or taking S7 to identify coordinate 0 of the dyad symmetry, they may be numbered -7 to +7.

Not all sites are cut with equal frequency: some are cut rather efficiently, others are cut scarcely at all. The enzymes DNAase I and DNAase II generate the same ladder, although with some differences in the intensities of the bands. This shows that the pattern of cutting represents a unique series of targets in DNA, determined by its organization, with only some slight preference for particular sites imposed by the individual enzyme. The same cutting pattern is obtained by cleaving with a hydroxyl radical, which argues that the pattern reflects the structure of the DNA itself, rather than any sequence preference.

The sensitivity of nucleosomal DNA to nucleases is analogous to a footprinting experiment. So we can assign the lack of reaction at particular target sites to the structure of the nucleosome, in which certain positions on DNA are rendered inaccessible.

Since there are two strands of DNA in the core particle, in an end-labeling experiment both 5′ (or both 3′) ends are labeled, one on each strand. So the cutting pattern includes fragments derived from both strands. This is implied in **Figure 20.11**, where each labeled fragment is derived from a different strand. The corollary is that, in an experiment, each labeled band in fact represents two fragments, generated by cutting the *same* distance from *either* of the labeled ends.

How then should we interpret discrete preferences at particular sites? One view is that the path of DNA on the particle is symmetrical (about a horizontal axis through the nucleosome drawn in **Figure 20.4**). So if (for example) no 80-base fragment is generated by DNAase I, this must mean that the position at 80 bases from the 5′ end of *either* strand is not susceptible to the enzyme. The second numbering scheme used in **Figure 20.13** reflects this view, and identifies S7 = site 0 as the center of symmetry.

When DNA is immobilized on a flat surface, sites are cut with a regular separation. **Figure 20.14** suggests that this reflects the recurrence of the exposed site with the helical periodicity of B-form DNA. The cutting periodicity (the spacing between cleavage points) coincides with, indeed, is a reflection of, the structural periodicity (the number of base pairs per turn of the double helix). So the distance between the sites corresponds to the number of base pairs per turn. Measurements of this type suggest that the average value for double-helical B-type DNA is 10.5 bp/turn.

Figure 20.14 The most exposed positions on DNA recur with a periodicity that reflects the structure of the double helix. (For clarity, sites are shown for only one strand.)

What is the nature of the target sites on the nucleosome? **Figure 20.15** shows that each site has 3-4 positions at which cutting occurs; that is, the cutting site is defined ±2 bp. So a cutting site represents a short stretch of bonds on both strands, exposed to nuclease action over 3-4 base pairs. The relative intensities indicate that some sites

are preferred to others.

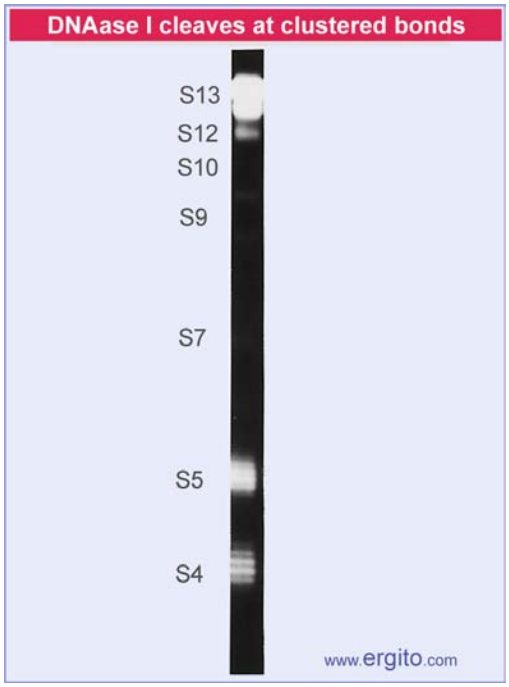

Figure 20.15 High resolution analysis shows that each site for DNAase I consists of several adjacent susceptible phosphodiester bonds as seen in this example of sites S4 and S5 analyzed in end-labeled core particles. Photograph kindly provided by Leonard Lutter.

From this pattern, we can calculate the "average" point that is cut. At the ends of the DNA, pairs of sites from S1 to S4 or from S10 to S13 lie apart a distance of 10.0 bases each. In the center of the particle, the separation from sites S4 to S10 is 10.7 bases. (Because this analysis deals with *average* positions, sites need not lie an integral number of bases apart.)

The variation in cutting periodicity along the core DNA (10.0 at the ends, 10.7 in the middle) means that there is variation in the structural periodicity of core DNA. The DNA has more bp/turn than its solution value in the middle, but has fewer bp/turn at the ends. The average periodicity over the nucleosome is only 10.17 bp/turn, which is significantly less than the 10.5 bp/turn of DNA in solution.

The crystal structure of the core particle suggests that DNA is organized as a flat superhelix, with 1.65 turns wound around the histone octamer. The pitch of the superhelix varies, with a discontinuity in the middle. Regions of high curvature are arranged symmetrically, and occur at positions ± 1 and ± 4. These correspond to S6 and S8, and to S3 and S11, which are the sites least sensitive to DNAase I (for review see 204).

A high resolution structure of the nucleosome core shows in detail how the structure of DNA is distorted (4003). Most of the supercoiling occurs in the central 129 bp, which are coiled into 1.59 left-handed superhelical turns with a diameter of 80

(only 4 ´ the diameter of the DNA duplex itself). The terminal sequences on either end make only a very small contribution to the overall curvature.

The central 129 bp are in the form of B-DNA, but with a substantial curvature that is needed to form the superhelix. The major groove is smoothly bent, but the minor groove has abrupt kinks. These conformational changes may explain why the central part of nucleosomal DNA is not usually a target for binding by regulatory proteins, which typically bind to the terminal parts of the core DNA or to the linker sequences.

Last updated on 7-21-2003

Reviews

204. Wang, J. (1982). *The path of DNA in the nucleosome*. Cell 29, 724-726.

References

4003. Richmond, T. J. and Davey, C. A. (2003). *The structure of DNA in the nucleosome core.* Nature 423, 145-150.

NUCLEOSOMES

5.20.6 The periodicity of DNA changes on the nucleosome

Key Concepts

- ~0.6 negative turns of DNA are absorbed by the change in bp/turn from 10.5 in solution to an average of 10.2 on the nucleosomal surface, explaining the linking number paradox.

Some insights into the structure of nucleosomal DNA emerge when we compare predictions for supercoiling in the path that DNA follows with actual measurements of supercoiling of nucleosomal DNA. Much work on the structure of sets of nucleosomes has been carried out with the virus SV40. The DNA of SV40 is a circular molecule of 5200 bp, with a contour length ~1500 nm. In both the virion and infected nucleus, it is packaged into a series of nucleosomes, called a minichromosome.

As usually isolated, the contour length of the minichromosome is ~210 nm, corresponding to a packing ratio of ~7 (essentially the same as the ~6 of the nucleosome itself). Changes in the salt concentration can convert it to a flexible string of beads with a much lower overall packing ratio. This emphasizes the point that nucleosome strings can take more than one form *in vitro*, depending on the conditions.

The degree of supercoiling on the individual nucleosomes of the minichromosome can be measured as illustrated in **Figure 20.16**. First, the free supercoils of the minichromosome itself are relaxed, so that the nucleosomes form a circular string with a superhelical density of 0. Then the histone octamers are extracted. This releases the DNA to follow a free path. Every supercoil that was present but restrained in the minichromosome will appear in the deproteinized DNA as –1 turn. So now the total number of supercoils in the SV40 DNA is measured.

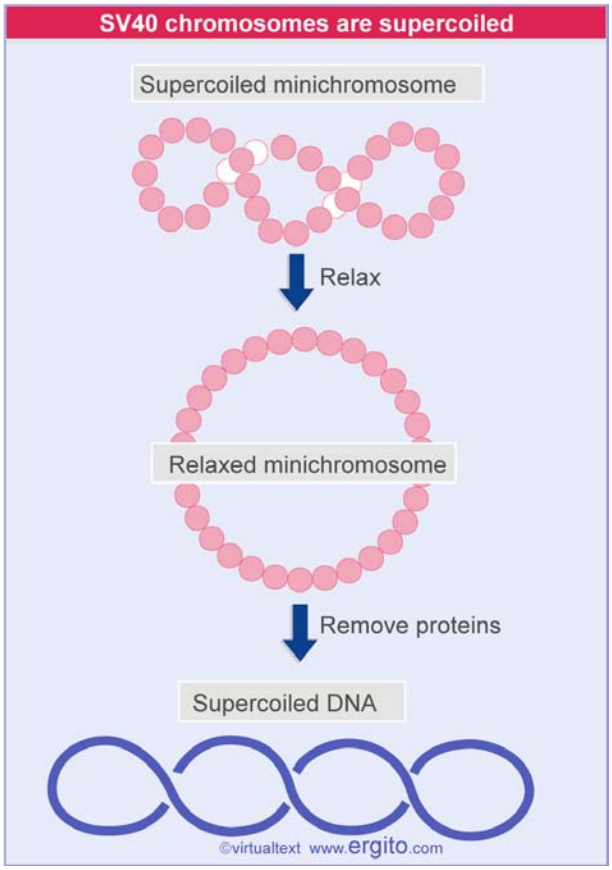

Figure 20.16 The supercoils of the SV40 minichromosome can be relaxed to generate a circular structure, whose loss of histones then generates supercoils in the free DNA.

The observed value is close to the number of nucleosomes. The reverse result is seen when nucleosomes are assembled *in vitro* on to a supercoiled SV40 DNA: the formation of each nucleosome removes ~1 negative supercoil.

So the DNA follows a path on the nucleosomal surface that generates ~1 negative supercoiled turn when the restraining protein is removed. But the path that DNA follows on the nucleosome corresponds to – 1.67 superhelical turns (see **Figure 20.4**). This discrepancy is sometimes called the linking number paradox .

The discrepancy is explained by the difference between the 10.17 average bp/turn of nucleosomal DNA and the 10.5 bp/turn of free DNA. In a nucleosome of 200 bp, there are 200/10.17 = 19.67 turns. When DNA is released from the nucleosome, it now has 200/10.5 = 19.0 turns. The path of the less tightly wound DNA on the nucleosome absorbs –0.67 turns, and this explains the discrepancy between the physical path of –1.67 and the measurement of –1.0 superhelical turns. In effect, some of the torsional strain in nucleosomal DNA goes into increasing the number of bp/turn; only the rest is left to be measured as a supercoil (for review see 206).

Last updated on 7-21-2003

Reviews

206. Travers, A. A. and Klug, A. (1987). *The bending of DNA in nucleosomes and its wider implications*. Philos. Trans. R. Soc. Lond. B Biol. Sci. 317, 537-561.

NUCLEOSOMES

5.20.7 The path of nucleosomes in the chromatin fiber

Key Terms

The **10 nm fiber** is a linear array of nucleosomes, generated by unfolding from the natural condition of chromatin.

The **30 nm fiber** is a coiled coil of nucleosomes. It is the basic level of organization of nucleosomes in chromatin.

Key Concepts

- 10 nm chromatin fibers are unfolded from 30 nm fibers and consist of a string of nucleosomes.
- 30 nm fibers have 6 nucleosomes/turn, organized into a solenoid.
- Histone H1 is required for formation of the 30 nm fiber.

When chromatin is examined in the electron microscope, two types of fibers are seen: the 10 nm fiber and 30 nm fiber. They are described by the approximate diameter of the thread (that of the 30 nm fiber actually varies from ~25–30 nm).

The **10 nm fiber** is essentially a continuous string of nucleosomes. Sometimes, indeed, it runs continuously into a more stretched-out region in which nucleosomes are seen as a string of beads, as indicated in the example of **Figure 20.17**. The 10 nm fibril structure is obtained under conditions of low ionic strength and does not require the presence of histone H1. This means that it is a function strictly of the nucleosomes themselves. It may be visualized essentially as a continuous series of nucleosomes, as in **Figure 20.18**. It is not clear whether such a structure exists *in vivo* or is simply a consequence of unfolding during extraction *in vitro*.

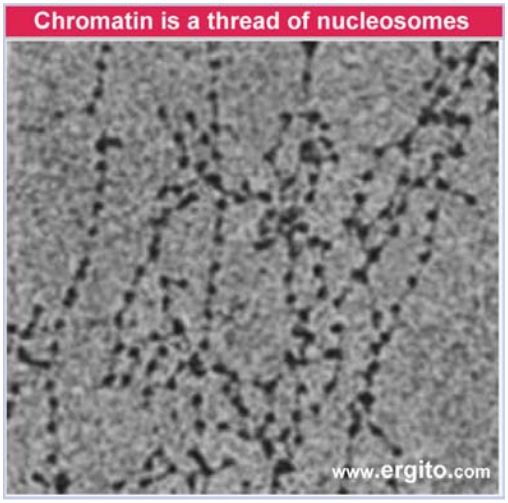

Figure 20.17 The 10 nm fiber in partially unwound state can be seen to consist of a string of nucleosomes. Photograph kindly provided by Barbara Hamkalo.

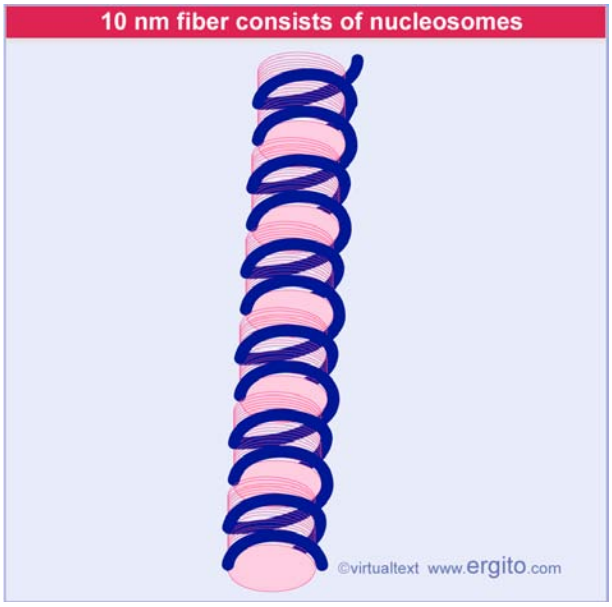

Figure 20.18 The 10 nm fiber is a continuous string of nucleosomes.

When chromatin is visualized in conditions of greater ionic strength the **30 nm fiber** is obtained. An example is given in **Figure 20.19**. The fiber can be seen to have an underlying coiled structure. It has ~6 nucleosomes for every turn, which corresponds to a packing ratio of 40 (that is, each µm along the axis of the fiber contains 40 µm of DNA). The presence of H1 is required. This fiber is the basic constituent of both interphase chromatin and mitotic chromosomes.

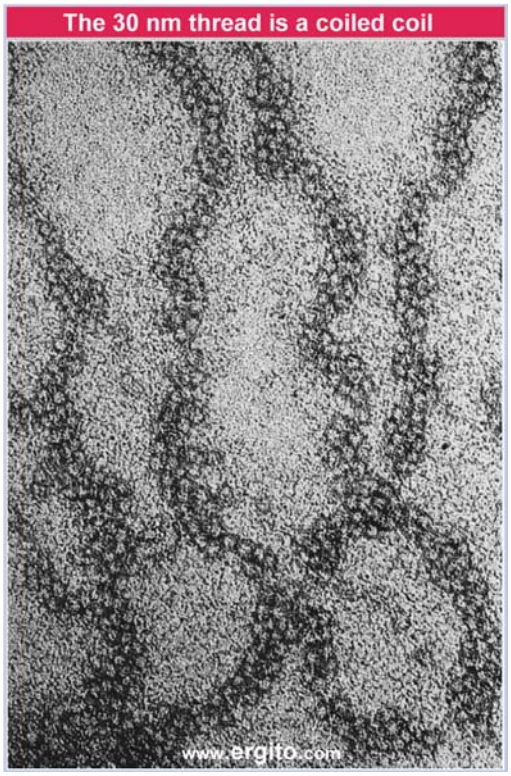

Figure 20.19 The 30 nm fiber has a coiled structure. Photograph kindly provided by Barbara Hamkalo.

The most likely arrangement for packing nucleosomes into the fiber is a solenoid, illustrated in **Figure 20.20**. The nucleosomes turn in a helical array, with an angle of ~60° between the faces of adjacent nucleosomes (for review see 205).

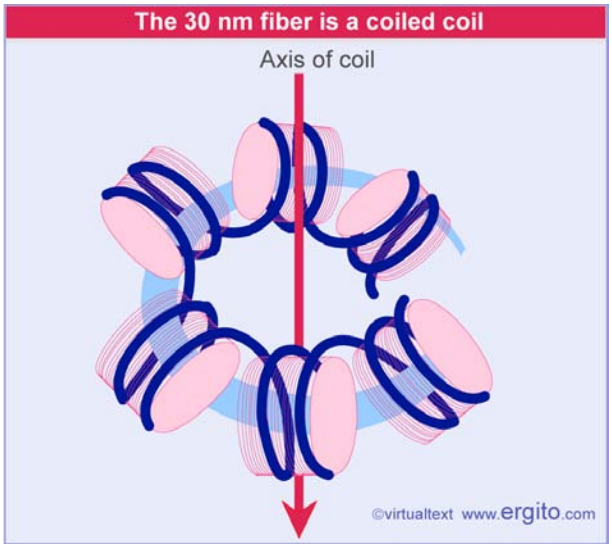

Figure 20.20 The 30 nm fiber may have a helical coil of 6 nucleosomes per turn, organized radially.

The 30 nm and 10 nm fibers can be reversibly converted by changing the ionic strength. This suggests that the linear array of nucleosomes in the 10 nm fiber is coiled into the 30 nm structure at higher ionic strength and in the presence of H1.

Although the presence of H1 is necessary for the formation of the 30 nm fiber, information about its location is conflicting. Its relative ease of extraction from chromatin seems to argue that it is present on the outside of the superhelical fiber axis. But diffraction data, and the fact that it is harder to find in 30 nm fibers than in 10 nm fibers that retain it, would argue for an interior location.

How do we get from the 30 nm fiber to the specific structures displayed in mitotic chromosomes? And is there any further specificity in the arrangement of interphase chromatin; do particular regions of 30 nm fibers bear a fixed relationship to one another or is their arrangement random?

Reviews

205. Felsenfeld, G. and McGhee, J. D. (1986). *Structure of the 30 nm chromatin fiber*. Cell 44, 375-377.

NUCLEOSOMES

5.20.8 Organization of the histone octamer

Key Terms

The **histone fold** is a motif found in all four core histones in which three α -helices are connected by two loops.

Key Concepts

- The histone octamer has a kernel of a $H3_2 \cdot H4_2$ tetramer associated with two H2A·H2B dimers.
- Each histone is extensively interdigitated with its partner.
- All core histones have the structural motif of the histone fold. N-terminal tails extend out of the nucleosome.

So far we have considered the construction of the nucleosome from the perspective of how the DNA is organized on the surface. From the perspective of protein, we need to know how the histones interact with each other and with DNA. Do histones react properly only in the presence of DNA, or do they possess an independent ability to form octamers? Most of the evidence about histone-histone interactions is provided by their abilities to form stable complexes, and by crosslinking experiments with the nucleosome.

The core histones form two types of complexes. H3 and H4 form a tetramer ($H3_2 \cdot H4_2$). Various complexes are formed by H2A and H2B, in particular a dimer (H2A·H2B).

Intact histone octamers can be obtained either by extraction from chromatin or (with more difficulty) by letting histones associate *in vitro* under conditions of high-salt and high-protein concentration. The octamer can dissociate to generate a hexamer of histones that has lost an H2A·H2B dimer. Then the other H2A·H2B dimer is lost separately, leaving the $H3_2 \cdot H4_2$ tetramer. This argues for a form of organization in which the nucleosome has a central "kernel" consisting of the $H3_2 \cdot H4_2$ tetramer. The tetramer can organize DNA *in vitro* into particles that display some of the properties of the core particle.

Crosslinking studies extend these relationships to show which pairs of histones lie near each other in the nucleosome. (A difficulty with such data is that usually only a small proportion of the proteins becomes crosslinked, so it is necessary to be cautious in deciding whether the results typify the major interactions.) From these data, a model has been constructed for the organization of the nucleosome. It is shown in diagrammatic form in **Figure 20.21**.

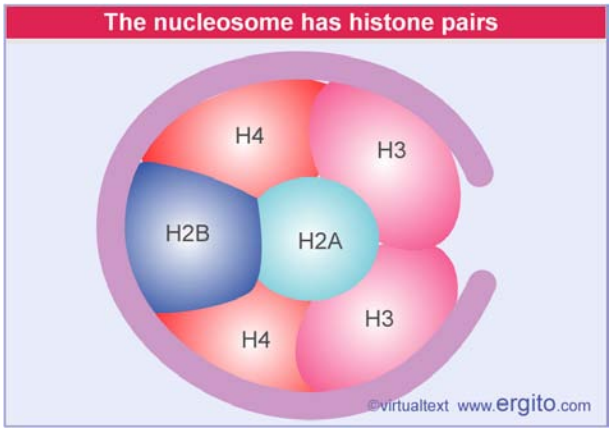

Figure 20.21 In a symmetrical model for the nucleosome, the $H3_2$- $H4_2$ tetramer provides a kernel for the shape. One H2A-H2B dimer can be seen in the top view; the other is underneath.

Structural studies show that the overall shape of the isolated histone octamer is similar to that of the core particle. This suggests that the histone-histone interactions establish the general structure. The positions of the individual histones have been assigned to regions of the octameric structure on the basis of their interaction behavior and response to crosslinking.

The crystal structure (at 3.1 Å resolution) suggests the model for the histone octamer shown in **Figure 20.22**. Tracing the paths of the individual polypeptide backbones in the crystal structure suggests that the histones are not organized as individual globular proteins, but that each is interdigitated with its partner, H3 with H4, and H2A with H2B. So the model distinguishes the $H3_2 \cdot H4_2$ tetramer (white) from the H2A·H2B dimers (blue), but does not show individual histones.

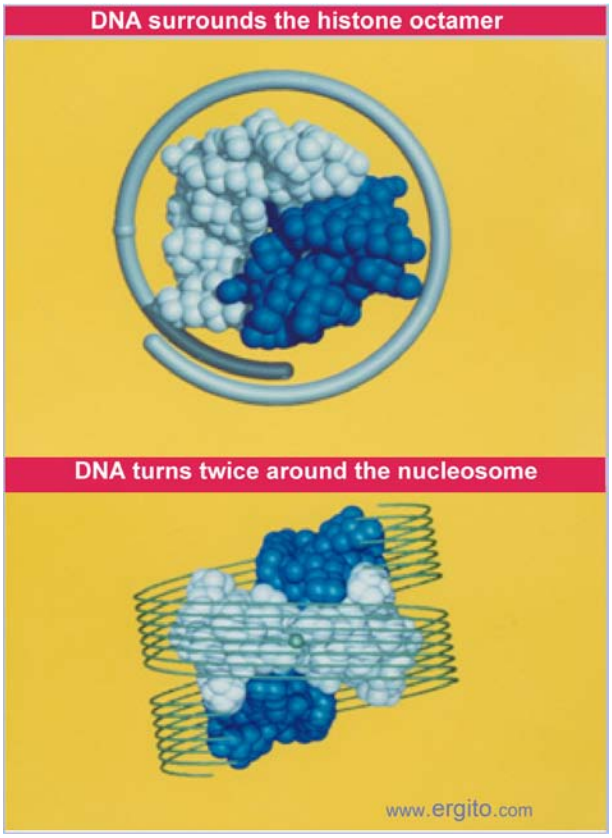

Figure 20.22 The crystal structure of the histone core octamer is represented in a space-filling model with the $H3_2$-$H4_2$ tetramer shown in white and the H2A-H2B dimers shown in blue. Only one of the H2A-H2B dimers is visible in the top view, because the other is hidden underneath. The potential path of the DNA is shown in the top view as a narrow tube (one quarter the diameter of DNA), and in the side view by the parallel lines in a 20 Å wide bundle. Photographs kindly provided by Evangelos Moudrianakis.

The top view represents the same perspective that was illustrated schematically in **Figure 20.21**. The $H3_2 \cdot H4_2$ tetramer accounts for the diameter of the octamer. It forms the shape of a horseshoe. The H2A·H2B pairs fit in as two dimers, but only one can be seen in this view. The side view represents the same perspective that was illustrated in **Figure 20.4**. Here the responsibilities of the $H3_2 \cdot H4_2$ tetramer and of the separate H2A·H2B dimers can be distinguished. The protein forms a sort of spool, with a superhelical path that could correspond to the binding site for DNA, which would be wound in almost two full turns in a nucleosome. The model displays two fold symmetry about an axis that would run perpendicular through the side view (601).

A more detailed view of the positions of the histones (based on a crystal structure at 2.8Å) is summarized in **Figure 20.23**. The upper view shows the position of one histone of each type relative to one turn around the nucleosome (numbered from 0 to +7). All four core histones show a similar type of structure in which three α-helices are connected by two loops: this is called the **histone fold**. These regions interact to form crescent-shaped heterodimers; each heterodimer binds 2.5 turns of the DNA

double helix (H2A-H2B binds at +3.5 - +6; H3-H4 binds at +0.5 - +3 for the circumference that is illustrated). Binding is mostly to the phosphodiester backbones (consistent with the need to package any DNA irrespective of sequence). The $H3_2 \cdot H4_2$ tetramer is formed by interactions between the two H3 subunits, as can be seen in the lower part of the figure (602).

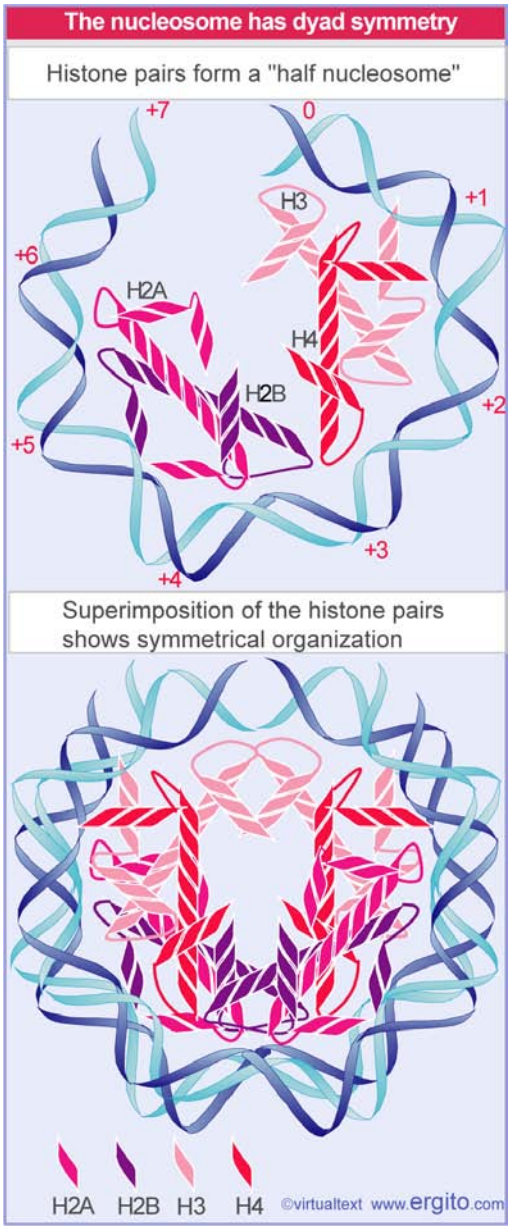

Figure 20.23 Histone positions in a top view show H3-H4 and H2A-H2B pairs in a half nucleosome; the symmetrical organization can be seen in the superimposition of both halves.

Each of the core histones has a globular body that contributes to the central protein mass of the nucleosome. Each histone also has a flexible N-terminal tail, which has sites for modification that may be important in chromatin function. The positions of

the tails, which account for about one quarter of the protein mass, are not so well defined, as indicated in **Figure 20.24**. However, the tails of both H3 and H2B can be seen to pass between the turns of the DNA superhelix and extend out of the nucleosome, as seen in **Figure 20.25**. When histone tails are crosslinked to DNA by UV irradiation, more products are obtained with nucleosomes compared to core particles, which could mean that the tails contact the linker DNA (2205). The tail of H4 appears to contact an H2A-H2B dimer in an adjacent nucleosome; this could be an important feature in the overall structure.

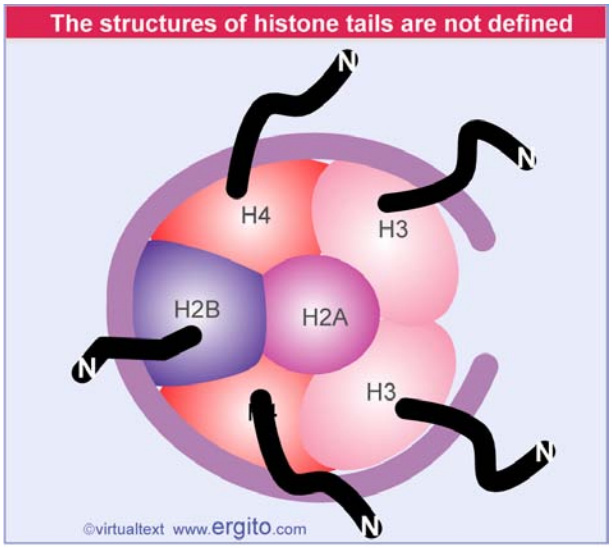

Figure 20.24 The globular bodies of the histones are localized in the histone octamer of the core particle, but the locations of the N-terminal tails, which carry the sites for modification, are not known, and could be more flexible.

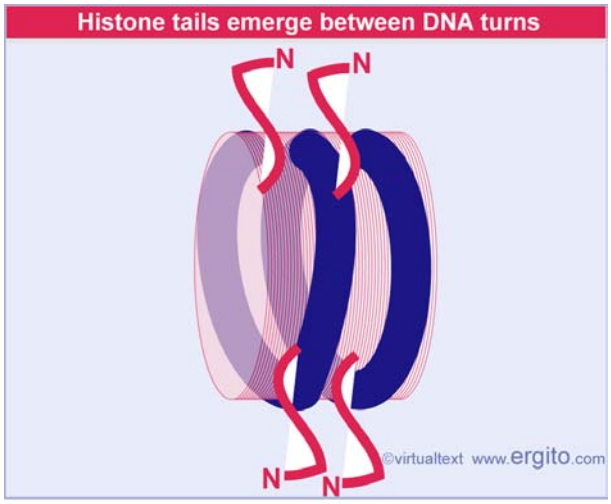

Figure 20.25 The N-terminal histone tails are disordered and exit from the nucleosome between turns of the DNA.

Last updated on 12-13-2001

References

601. Arents, G., Burlingame, R. W., Wang, B.-C., Love, W. E., and Moudrianakis, E. N. (1991). *The nucleosomal core histone octamer at 31* Å *resolution: a tripartite protein assembly and a left-handed superhelix*. Proc. Natl. Acad. Sci. USA 88, 10148-10152.

602. Luger, K. et al. (1997). *Crystal structure of the nucleosome core particle at 28* Å *resolution*. Nature 389, 251-260.

2205. Angelov, D., Vitolo, J. M., Mutskov, V., Dimitrov, S., and Hayes, J. J. (2001). *Preferential interaction of the core histone tail domains with linker DNA.* Proc. Natl. Acad. Sci. USA 98, 6599-6604.

NUCLEOSOMES

5.20.9 The N-terminal tails of histones are modified

Key Concepts

- Histones are modified by methylation, acetylation, and phosphorylation.

All of the histones are modified by covalently linking extra moieties to the free groups of certain amino acids. The sites that are modified are concentrated in the N-terminal tails. These modifications have important effects on the structure of chromatin and in controlling gene expression (see *Molecular Biology 5.23.5 Histone modification is a key event*).

Acetylation and methylation occur on the free (ε) amino group of lysine. As seen in **Figure 20.26**, this removes the positive charge that resides on the NH^+_3 form of the group. Methylation also occurs on arginine and histidine. Phosphorylation occurs on the hydroxyl group of serine and also on histidine. This introduces a negative charge in the form of the phosphate group.

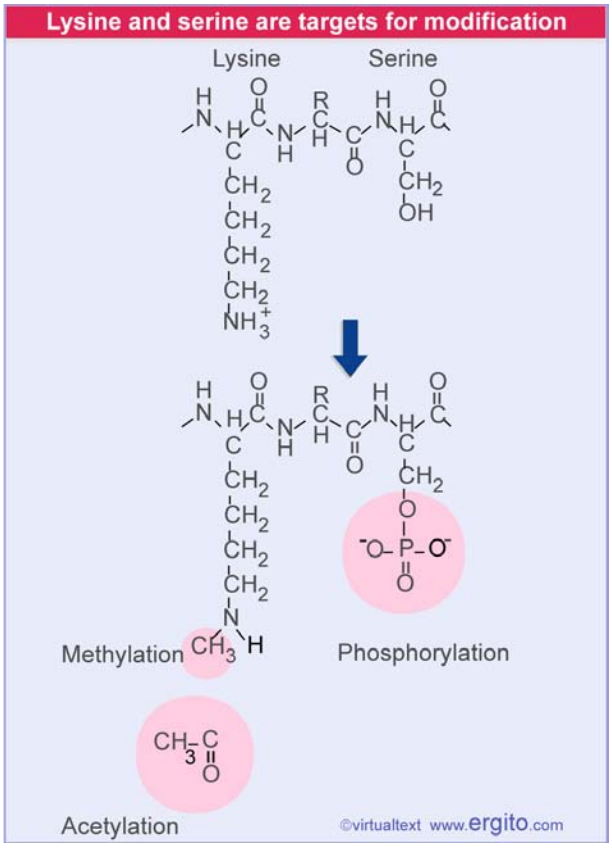

Figure 20.26 Acetylation of lysine or phosphorylation of serine reduces the overall positive charge of a protein.

These modifications are transient. Because they change the charge of the protein molecule, they are potentially able to change the functional properties of the octamers. Modification of histones is associated with structural changes that occur in chromatin at replication and transcription. Phosphorylations on specific positions and on different histones may be required for particular processes, for example, the Ser^{10} position of H3 is phosphorylated when chromosomes condense at mitosis.

In synchronized cells in culture, both the pre-existing and newly synthesized core histones appear to be acetylated and methylated during S phase (when DNA is replicated and the histones also are synthesized). During the cell cycle, the modifying groups are later removed.

The coincidence of modification and replication suggests that acetylation (and methylation) could be connected with nucleosome assembly. One speculation has been that the reduction of positive charges on histones might lower their affinity for DNA, allowing the reaction to be better controlled. The idea has lost some ground in view of the observation that nucleosomes can be reconstituted, at least *in vitro,* with unmodified histones. Histone acetylation is essential for nucleosome assembly in yeast, and is probably required for some of the protein-protein interactions that occur during later stages of the reaction (see *Molecular Biology 5.23.6 Histone acetylation occurs in two circumstances*).

A cycle of phosphorylation and dephosphorylation occurs with H1, but its timing is different from the modification cycle of the other histones. With cultured mammalian cells, one or two phosphate groups are introduced at S phase. But the major phosphorylation event is the later addition of more groups at mitosis, to bring the total number up to as many as six. All the phosphate groups are removed at the end of the process of division. The phosphorylation of H1 is catalyzed by the M-phase kinase that provides an essential trigger for mitosis (see *Molecular Biology 6.29 Cell cycle and growth regulation*). In fact, this enzyme is now often assayed in terms of its H1 kinase activity. Not much is known about phosphatase(s) that remove the groups later.

The timing of the major H1 phosphorylation has prompted speculation that it is involved in mitotic condensation. However, in *Tetrahymena* (a protozoan) it is possible to delete all the genes for H1 without significantly affecting the overall properties of chromatin. There is a relatively small effect on the ability of chromatin to condense at mitosis. Some genes are activated and others are repressed by this change, suggesting that there alterations in local structure. Mutations that eliminate sites of phosphorylation in H1 have no effect, but mutations that mimic the effects of phosphorylation produce a phenotype that resembles the deletion. This suggests that the effect of phosphorylating H1 is to eliminate its effects on local chromatin structure.

NUCLEOSOMES

5.20.10 Reproduction of chromatin requires assembly of nucleosomes

Key Concepts

- Histone octamers are not conserved during replication, but H2A·H2B dimers and $H3_2 \cdot H4_2$ tetramers are conserved.
- There are different pathways for the assembly of nucleosomes during replication and independently of replication.
- Accessory proteins are required to assist the assembly of nucleosomes.
- CAF-1 is an assembly protein that is linked to the PCNA subunit of the replisome; it is required for deposition of $H3_2 \cdot H4_2$ tetramers following replication.
- A different assembly protein and a variant of histone H3 may be used for replication-independent assembly.

Replication separates the strands of DNA and therefore must inevitably disrupt the structure of the nucleosome. The transience of the replication event is a major difficulty in analyzing the structure of a particular region while it is being replicated. The structure of the replication fork is distinctive. It is more resistant to micrococcal nuclease and is digested into bands that differ in size from nucleosomal DNA. The region that shows this altered structure is confined to the immediate vicinity of the replication fork. This suggests that a large protein complex is engaged in replicating the DNA, but the nucleosomes reform more or less immediately behind as it moves along.

Reproduction of chromatin does not involve any protracted period during which the DNA is free of histones. Once DNA has been replicated, nucleosomes are quickly generated on both the duplicates. This point is illustrated by the electron micrograph of **Figure 20.27**, which shows a recently replicated stretch of DNA, already packaged into nucleosomes on both daughter duplex segments.

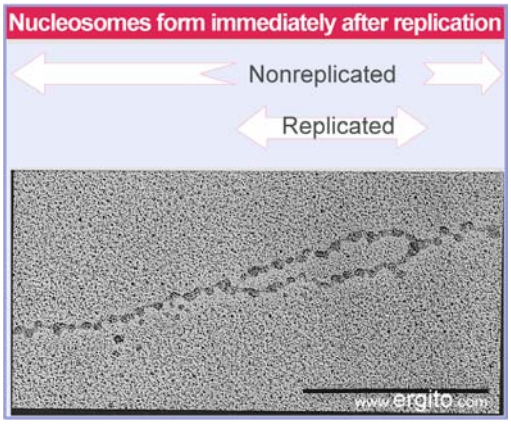

Figure 20.27 Replicated DNA is immediately incorporated into nucleosomes. Photograph kindly provided by S. MacKnight.

Both biochemical analysis and visualization of the replication fork therefore suggest that the disruption of nucleosome structure is limited to a short region immediately around the fork. Progress of the fork disrupts nucleosomes, but they form very rapidly on the daughter duplexes as the fork moves forward. In fact, the assembly of nucleosomes is directly linked to the replisome that is replicating DNA (for review see 1978).

How do histones associate with DNA to generate nucleosomes? Do the histones *preform* a protein octamer around which the DNA is subsequently wrapped? Or does the histone octamer assemble on DNA from free histones? **Figure 20.28** shows that two pathways can be used *in vitro* to assemble nucleosomes, depending on the conditions that are employed. In one pathway, a preformed octamer binds to DNA. In the other pathway, a tetramer of $H3_2 \cdot H4_2$ binds first, and then two H2A·H2B dimers are added. Both these pathways are related to reactions that occur *in vivo*. The first reflects the capacity of chromatin to be remodeled by moving histone octamers along DNA (see *Molecular Biology 5.23.3 Chromatin remodeling is an active process*). The second represents the pathway that is used in replication.

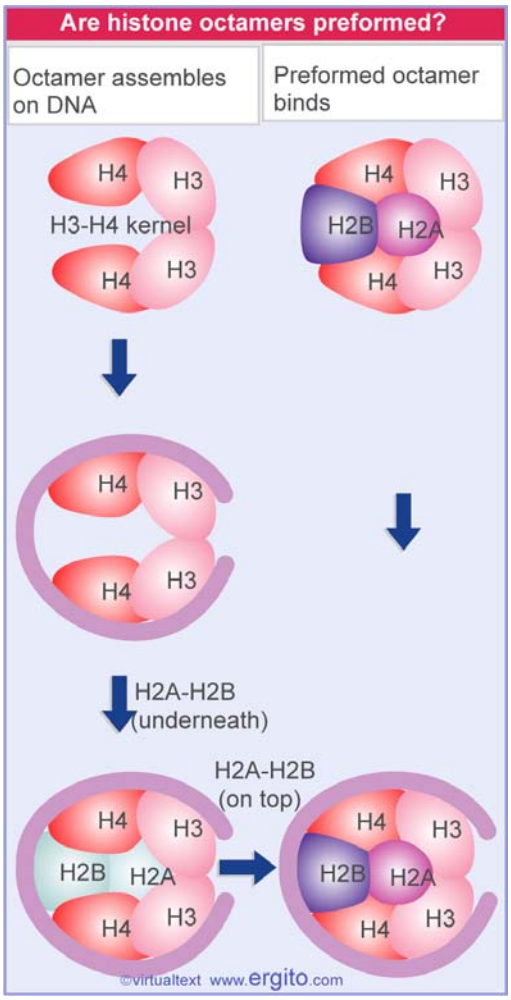

Figure 20.28 *In vitro*, DNA can either interact directly with an intact (crosslinked) histone octamer or can assemble with the $H3_2$-$H4_2$ tetramer, after which two H2A-H2B dimers are added.

Accessory proteins are involved in assisting histones to associate with DNA. Candidates for this role can be identified by using extracts that assemble histones and exogenous DNA into nucleosomes. Accessory proteins may act as "molecular chaperones" that bind to the histones in order to release either individual histones or complexes ($H3_2 \cdot H4_2$ or H2A·H2B) to the DNA in a controlled manner. This could be necessary because the histones, as basic proteins, have a general high affinity for DNA. *Such interactions allow histones to form nucleosomes without becoming trapped in other kinetic intermediates (that is, other complexes resulting from indiscreet binding of histones to DNA).*

Attempts to produce nucleosomes *in vitro* began by considering a process of assembly between free DNA and histones. But nucleosomes form *in vivo* only when DNA is replicated. A system that mimics this requirement has been developed by using extracts of human cells that replicate SV40 DNA and assemble the products into chromatin. The assembly reaction occurs preferentially on replicating DNA. It requires an ancillary factor, CAF-1, that consists of >5 subunits, with a total mass of

238 kD (3198; 3199). CAF-1 is recruited to the replication fork by PCNA, the processivity factor for DNA polymerase. This provides the link between replication and nucleosome assembly, ensuring that nucleosomes are assembled as soon as DNA has been replicated (609).

CAF-1 acts stoichiometrically, and functions by binding to newly synthesized H3 and H4. This suggests that new nucleosomes form by assembling first the $H3_2 \cdot H4_2$ tetramer, and then adding the H2A·H2B dimers. The nucleosomes that are formed *in vitro* have a repeat length of 200 bp, although they do not have any H1 histone, which suggests that proper spacing can be accomplished without H1.

When chromatin is reproduced, a stretch of DNA *already associated with nucleosomes* is replicated, giving rise to two daughter duplexes. What happens to the pre-existing nucleosomes at this point? Are the histone octamers dissociated into free histones for reuse, or do they remain assembled? The integrity of the octamer can be tested by crosslinking the histones. The next two figures compare the possible outcomes from an experiment in which cells are grown in the presence of heavy amino acids to identify the histones before replication. Then replication is allowed to occur in the presence of light amino acids. At this point the histone octamers are crosslinked and centrifuged on a density gradient. **Figure 20.29** shows that if the original octamers have been conserved, they will be found at a position of high density, and new octamers will occupy a low density position. However, this does not happen. Little material is found at the high density position, which suggests that histone octamers are not conserved. The octamers have an intermediate density, and **Figure 20.30** shows that this is the expected result if the old histones have been released and then reassembled with newly synthesized histones.

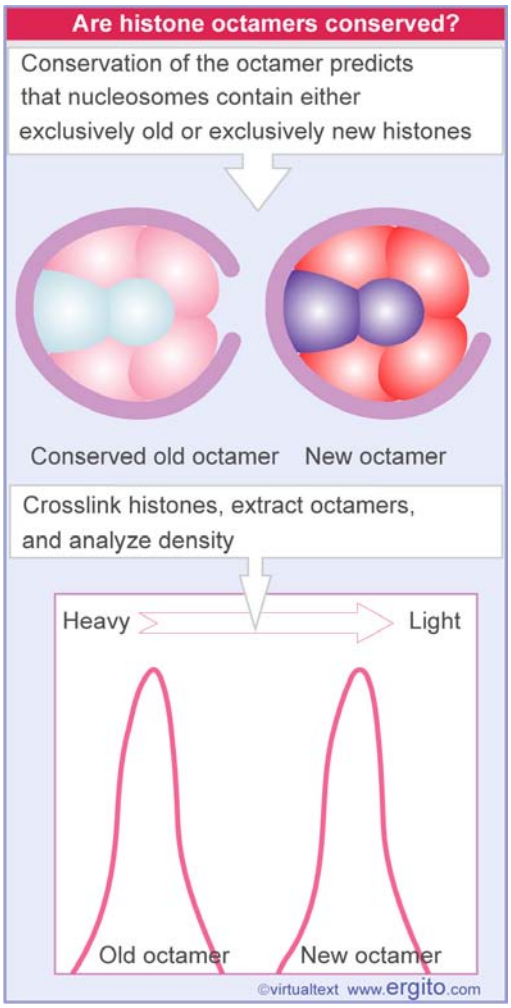

Figure 20.29 If histone octamers were conserved, old and new octamers would band at different densities when replication of heavy octamers occurs in light amino acids.

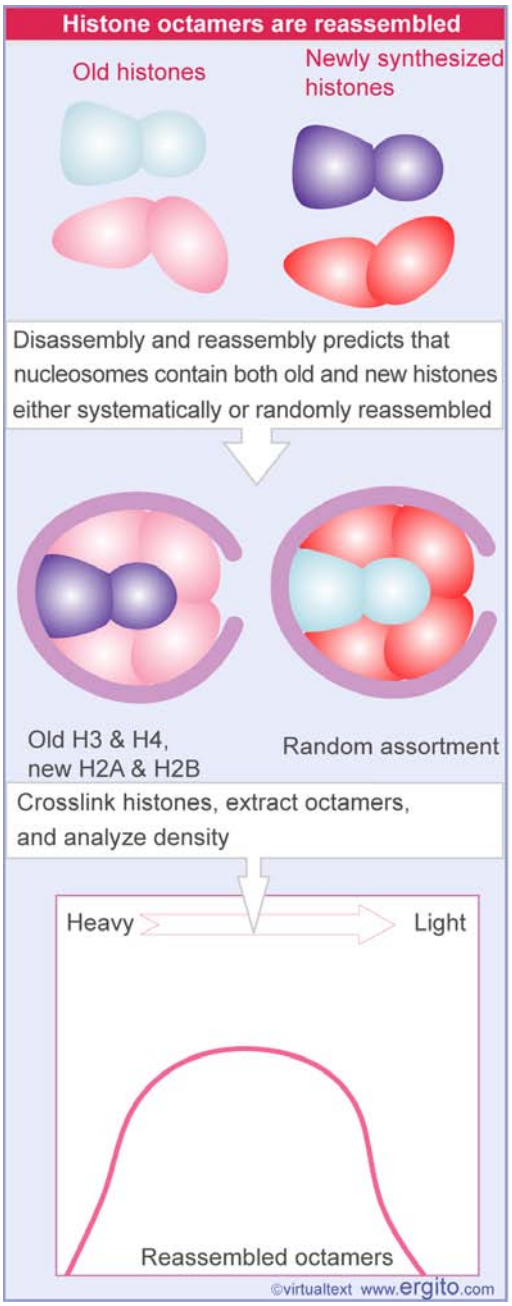

Figure 20.30 When heavy octamers are replicated in light amino acids, the new octamers band diffusely between heavy and light densities, suggesting that disassembly and reassembly has occurred.

The pattern of disassembly and reassembly has been difficult to characterize in detail, but our working model is illustrated in **Figure 20.31**. The replication fork displaces histone octamers, which then dissociate into $H3_2 \cdot H4_2$ tetramers and H2A·H2B dimers (3197). These "old" tetramers and dimers enter a pool that also includes "new" tetramers and dimers, assembled from newly synthesized histones. Nucleosomes assemble ~600 bp behind the replication fork. Assembly is initiated when $H3_2 \cdot H4_2$ tetramers bind to each of the daughter duplexes, assisted by CAF-1.

Then two H2A·H2B dimers bind to each $H3_2 \cdot H4_2$ tetramer to complete the histone octamer. The assembly of tetramers and dimers is random with respect to "old" and "new" subunits, explaining the results of **Figure 20.30**. The "old" $H3_2 \cdot H4_2$ tetramer could have an ability to be transiently associated with a single strand of DNA during replication; it may in fact have an increased chance of remaining with the leading strand for reuse. It is possible that nucleosomes are disrupted and reassembled in a similar way during transcription (see *Molecular Biology 5.20.12 Are transcribed genes organized in nucleosomes?*).

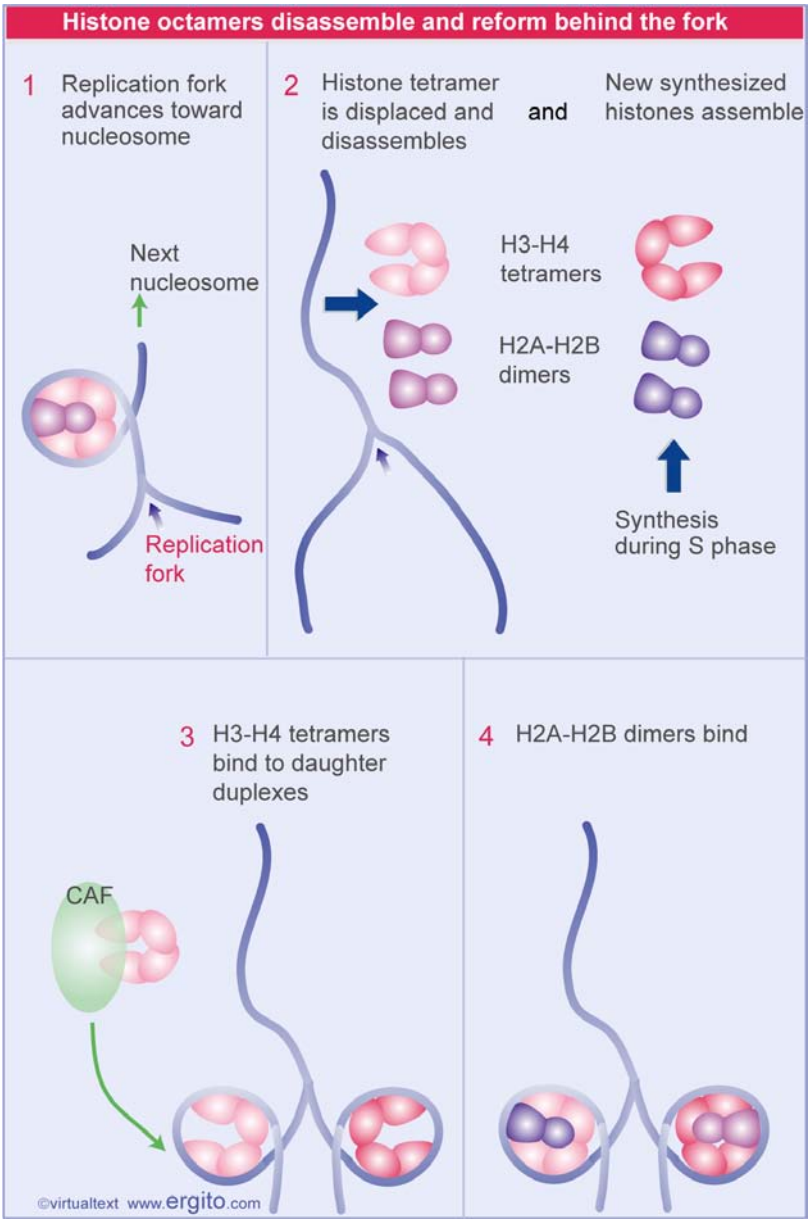

Figure 20.31 Replication fork passage displaces histone octamers from DNA. They disassemble into H3-H4 tetramers and H2A-H2B dimers. Newly synthesized histones are assembled into H3-H4 tetramers and H2A-H2B dimers. The old and new tetramers and dimers are assembled with the aid of CAF-1 at random into new nucleosomes immediately behind the replication fork

During S phase (the period of DNA replication) in a eukaryotic cell, the duplication of chromatin requires synthesis of sufficient histone proteins to package an entire genome – basically the same quantity of histones must be synthesized that are already contained in nucleosomes. The synthesis of histone mRNAs is controlled as part of the cell cycle, and increases enormously in S phase (for review see 3200). The pathway for assembling chromatin from this equal mix of old and new histones during S phase is called the replication-coupled (RC) pathway.

Another pathway, called the replication-independent (RI) pathway exists for assembling nucleosomes during other phases of cell cycle, when DNA is not being synthesized. This may become necessary as the result of damage to DNA or because nucleosomes are displaced during transcription. The assembly process must necessarily have some differences from the replication-coupled pathway, because it cannot be linked to the replication apparatus. One of the most interesting features of the replication-independent pathway is that it uses different variants of some of the histones from those used during replication.

The histone H3.3 variant differs from the highly conserved H3 histone at 4 amino acid positions. H3.3 slowly replaces H3 in differentiating cells that do not have replication cycles. This happens as the result of assembly of new histone octamers to replace those that have been displaced from DNA for whatever reason. The mechanism that is used to ensure the use of H3.3 in the replication-independent pathway is different in two cases that have been investigated.

In the protozoan *Tetrahymena*, histone usage is determined exclusively by availability. Histone H3 is synthesized only during the cell cycle; the variant replacement histone is synthesized only in nonreplicating cells (3201). In *Drosophila*, however, there is an active pathway that ensures the usage of H3.3 by the replication-independent pathway. New nucleosomes containing H3.3 assemble at sites of transcription, presumably replacing nucleosome that were displaced by RNA polymerase. The assembly process discriminates between H3 and H3.3 on the basis of their sequences, specifically excluding H3 from being utilized (3185). By contrast, replication-coupled assembly uses both types of H3 (although H3.3 is available at much lower levels than H3, and therefore enters only a small proportion of nucleosomes).

CAF-1 is probably not involved in replication-independent assembly. (And there are organisms such as yeast and *Arabidopsis* where its gene is not essential, implying that alternative assembly processes may be used in replication-coupled assembly). A protein that may be involved in replication-independent assembly is called HIRA. Depletion of HIRA from *in vitro* systems for nucleosome assembly inhibits the formation of nucleosomes on nonreplicated DNA, but not on replicating DNA, indicating that the pathways do indeed use different assembly mechanisms (3190).

Assembly of nucleosomes containing an alternative to H3 also occurs at centromeres (see *Molecular Biology 5.23.15 Heterochromatin depends on interactions with histones*). Centromeric DNA replicates early during the replication phase of the cell cycle (in contrast with the surrounding heterochromatic sequences that replicate later; see *Molecular Biology 4.13.5 Each eukaryotic chromosome contains many replicons*). The incorporation of H3 at the centromeres is inhibited, and instead a protein called CENP-A is incorporated in higher eukaryotic cells (in *Drosophila* it is called Cid, and in yeast it is called Cse4p). This occurs by the

replication-independent assembly pathway, apparently because the replication-coupled pathway is inhibited for a brief period of time while centromeric DNA replicates (3188).

Last updated on 12-16-2002

Reviews

1978. Verreault, A. (2000). *De novo nucleosome assembly: new pieces in an old puzzle.* Genes Dev. 14, 1430-1438.

3200. Osley, M. A. (1991). *The regulation of histone synthesis in the cell cycle.* Annu. Rev. Biochem. 60, 827-861.

References

609. Shibahara, K., and Stillman, B. (1999). *Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin.* Cell 96, 575-585.

3185. Ahmad, K. and Henikoff, S. (2002). *The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly.* Mol. Cell 9, 1191-1200.

3188. Ahmad, K. and Henikoff, S. (2001). *Centromeres are specialized replication domains in heterochromatin.* J. Cell Biol. 153, 101-110.

3190. Ray-Gallet, D., Quivy, J. P., Scamps, C., Martini, E. M., Lipinski, M., and Almouzni, G. (2002). *HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis.* Mol. Cell 9, 1091-1100.

3197. Gruss, C., Wu, J., Koller, T., and Sogo, J. M. (1993). *Disruption of the nucleosomes at the replication fork.* EMBO J. 12, 4533-4545.

3198. Smith, S. and Stillman, B. (1989). *Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro.* Cell 58, 15-25.

3199. Smith, S. and Stillman, B. (1991). *Stepwise assembly of chromatin during DNA replication in vitro.* EMBO J. 10, 971-980.

3201. Yu, L. and Gorovsky, M. A. (1997). *Constitutive expression, not a particular primary sequence, is the important feature of the H3 replacement variant hv2 in Tetrahymena thermophila.* Mol. Cell. Biol. 17, 6303-6310.

NUCLEOSOMES

5.20.11 Do nucleosomes lie at specific positions?

Key Terms

Nucleosome positioning describes the placement of nucleosomes at defined sequences of DNA instead of at random locations with regards to sequence.

Indirect end labeling is a technique for examining the organization of DNA by making a cut at a specific site and isolating all fragments containing the sequence adjacent to one side of the cut; it reveals the distance from the cut to the next break(s) in DNA.

Translational positioning describes the location of a histone octamer at successive turns of the double helix, which determines which sequences are located in linker regions.

Rotational positioning describes the location of the histone octamer relative to turns of the double helix, which determines which face of DNA is exposed on the nucleosome surface.

Key Concepts

- Nucleosomes may form at specific positions as the result either of the local structure of DNA or of proteins that interact with specific sequences.
- The most common cause of nucleosome positioning is when proteins binding to DNA establish a boundary.
- Positioning may affect which regions of DNA are in the linker and which face of DNA is exposed on the nucleosome surface.

We know that nucleosomes can be reconstituted *in vitro* without regard to DNA sequence, but this does not mean that their formation *in vivo* is independent of sequence. Does a particular DNA sequence always lie in a certain position *in vivo* with regard to the topography of the nucleosome? Or are nucleosomes arranged randomly on DNA, so that a particular sequence may occur at any location, for example, in the core region in one copy of the genome and in the linker region in another?

To investigate this question, it is necessary to use a defined sequence of DNA; more precisely, we need to determine the position relative to the nucleosome of a defined point in the DNA. **Figure 20.32** illustrates the principle of a procedure used to achieve this.

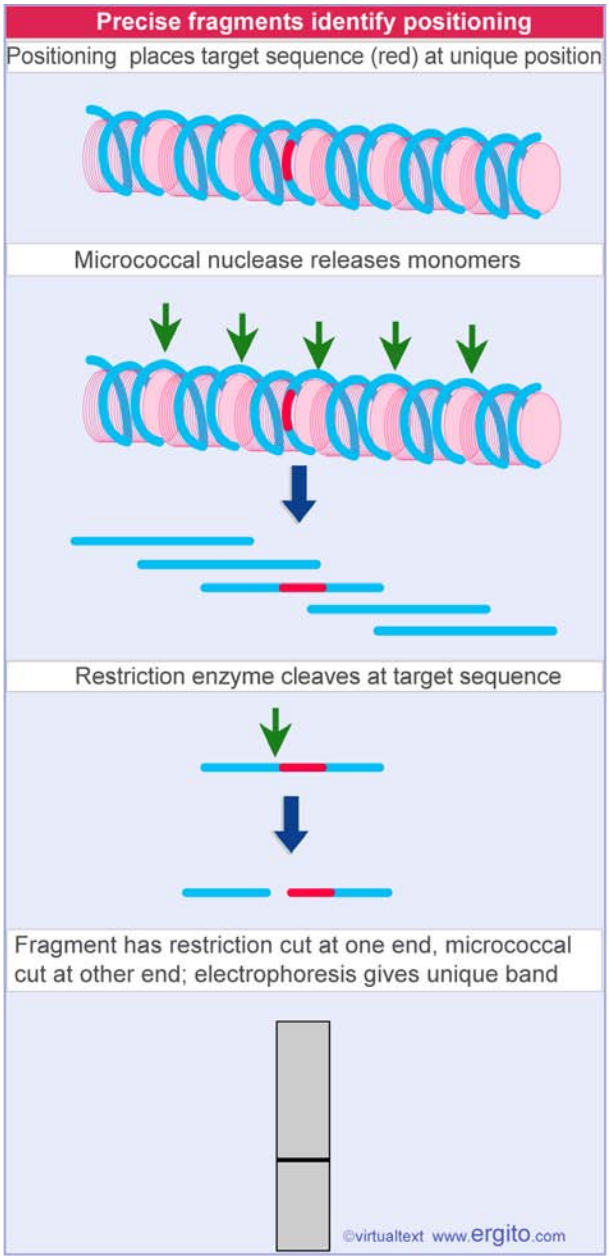

Figure 20.32 Nucleosome positioning places restriction sites at unique positions relative to the linker sites cleaved by micrococcal nuclease.

Suppose that the DNA sequence is organized into nucleosomes in only one particular configuration, so that each site on the DNA always is located at a particular position on the nucleosome. This type of organization is called **nucleosome positioning** (or sometimes nucleosome phasing). In a series of positioned nucleosomes, the linker regions of DNA comprise unique sites.

Consider the consequences for just a single nucleosome. Cleavage with micrococcal nuclease generates a monomeric fragment that constitutes a *specific sequence*. If the DNA is isolated and cleaved with a restriction enzyme that has only one target site in

this fragment, it should be cut at a unique point. This produces two fragments, each of unique size.

The products of the micrococcal/restriction double digest are separated by gel electrophoresis. A probe representing the sequence on one side of the restriction site is used to identify the corresponding fragment in the double digest. This technique is called **indirect end labeling**.

Reversing the argument, the identification of a single sharp band demonstrates that the position of the restriction site is uniquely defined with respect to the end of the nucleosomal DNA (as defined by the micrococcal nuclease cut). So the nucleosome has a unique sequence of DNA.

What happens if the nucleosomes do *not* lie at a single position? Now the linkers consist of *different* DNA sequences in each copy of the genome. So the restriction site lies at a different position each time; in fact, it lies at all possible locations relative to the ends of the monomeric nucleosomal DNA. **Figure 20.33** shows that the double cleavage then generates a broad smear, ranging from the smallest detectable fragment (~20 bases) to the length of the monomeric DNA.

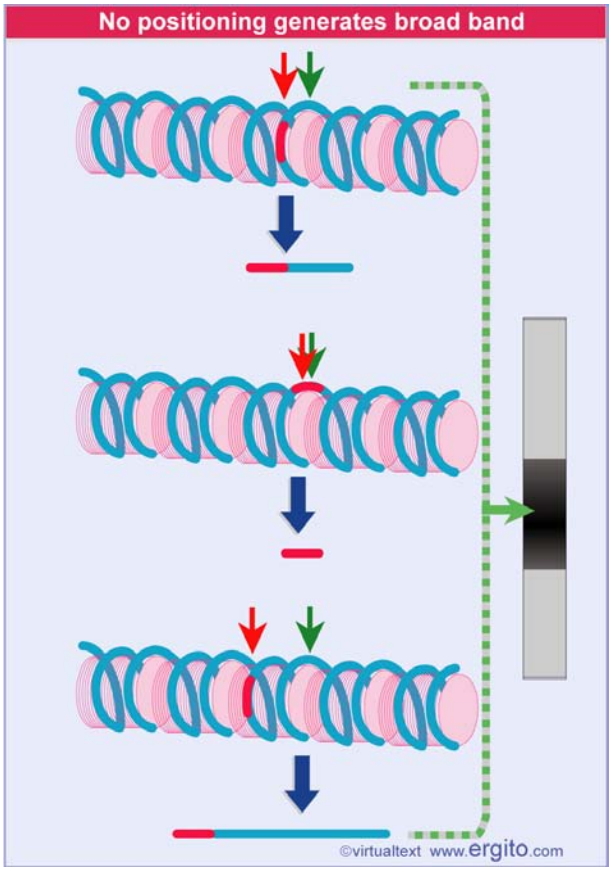

Figure 20.33 In the absence of nucleosome positioning, a restriction site lies at all possible locations in different copies of the genome. Fragments of all possible sizes are produced when a restriction enzyme cuts at a target site (red) and micrococcal nuclease cuts at the junctions between nucleosomes (green).

In discussing these experiments, we have treated micrococcal nuclease as an enzyme that cleaves DNA at the exposed linker regions without any sort of sequence specificity. However, the enzyme actually does have some sequence specificity (biased toward selection of A·T-rich sequences). So we cannot assume that the existence of a specific band in the indirect end-labeling technique represents the distance from a restriction cut to the linker region. It could instead represent the distance from the restriction cut to a preferred micrococcal nuclease cleavage site!

This possibility is controlled by treating the naked DNA in exactly the same way as the chromatin. If there are preferred sites for micrococcal nuclease in the particular region, specific bands are found. Then this pattern of bands can be compared with the pattern generated from chromatin.

A *difference* between the control DNA band pattern and the chromatin pattern provides evidence for nucleosome positioning. Some of the bands present in the control DNA digest may disappear from the nucleosome digest, indicating that preferentially cleaved positions are unavailable. New bands may appear in the nucleosome digest when new sites are rendered preferentially accessible by the nucleosomal organization.

Nucleosome positioning might be accomplished in either of two ways:

- It is intrinsic: *every nucleosome is deposited specifically at a particular DNA sequence.* This modifies our view of the nucleosome as a subunit able to form between any sequence of DNA and a histone octamer.
- It is extrinsic: *the first nucleosome in a region is preferentially assembled at a particular site.* A preferential starting point for nucleosome positioning results from the presence of a region from which nucleosomes are excluded. The excluded region provides a *boundary* that restricts the positions available to the adjacent nucleosome. Then a series of nucleosomes may be assembled sequentially, with a defined repeat length.

It is now clear that the deposition of histone octamers on DNA is not random with regard to sequence. The pattern is intrinsic in some cases, in which it is determined by structural features in DNA. It is extrinsic in other cases, in which it results from the interactions of other proteins with the DNA and/or histones.

Certain structural features of DNA affect placement of histone octamers. DNA has intrinsic tendencies to bend in one direction rather than another; thus A·T-rich regions locate so that the minor groove faces in towards the octamer, whereas G·C-rich regions are arranged so that the minor groove points out. Long runs of dA·dT (>8 bp) avoid positioning in the central superhelical turn of the core. It is not yet possible to sum all of the relevant structural effects and thus entirely to predict the location of a particular DNA sequence with regard to the nucleosome. Sequences that cause DNA to take up more extreme structures may have effects such as the exclusion of nucleosomes, and thus could cause boundary effects.

Positioning of nucleosomes near boundaries is common. If there is some variability in the construction of nucleosomes – for example, if the length of the linker can vary

by, say, 10 bp – the specificity of location would decline proceeding away from the first, defined nucleosome at the boundary. In this case, we might expect the positioning to be maintained rigorously only relatively near the boundary.

The location of DNA on nucleosomes can be described in two ways. **Figure 20.34** shows that **translational positioning** describes the position of DNA with regard to the boundaries of the nucleosome. In particular, it determines which sequences are found in the linker regions. Shifting the DNA by 10 bp brings the next turn into a linker region. So translational positioning determines which regions are more accessible (at least as judged by sensitivity to micrococcal nuclease).

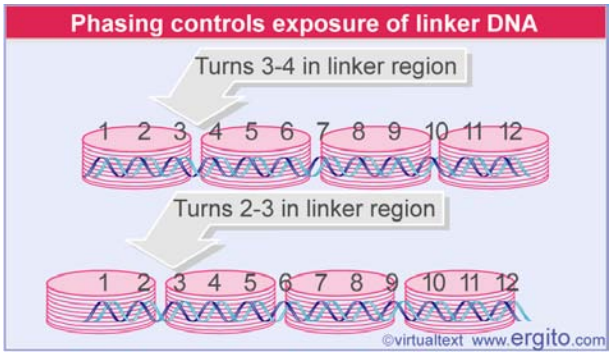

Figure 20.34 Translational positioning describes the linear position of DNA relative to the histone octamer. Displacement of the DNA by 10 bp changes the sequences that are in the more exposed linker regions, but does not alter which face of DNA is protected by the histone surface and which is exposed to the exterior. DNA is really coiled around the nucleosomes, and is shown in linear form only for convenience.

Because DNA lies on the outside of the histone octamer, one face of any particular sequence is obscured by the histones, but the other face is accessible. Depending upon its positioning with regard to the nucleosome, a site in DNA that must be recognized by a regulator protein could be inaccessible or available. The exact position of the histone octamer with respect to DNA sequence may therefore be important. **Figure 20.35** shows the effect of **rotational positioning** of the double helix with regard to the octamer surface. If the DNA is moved by a partial number of turns (imagine the DNA as rotating relative to the protein surface), there is a change in the exposure of sequence to the outside.

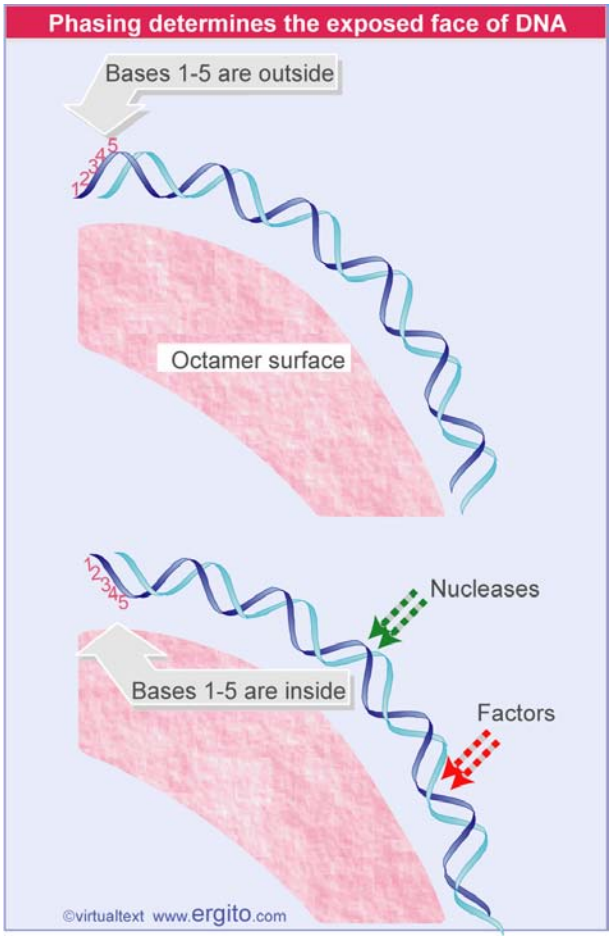

Figure 20.35 Rotational positioning describes the exposure of DNA on the surface of the nucleosome. Any movement that differs from the helical repeat (~10.2 bp/turn) displaces DNA with reference to the histone surface. Nucleotides on the inside are more protected against nucleases than nucleotides on the outside.

Both translational and rotational positioning can be important in controlling access to DNA. The best characterized cases of positioning involve the specific placement of nucleosomes at promoters. Translational positioning and/or the exclusion of nucleosomes from a particular sequence may be necessary to allow a transcription complex to form. Some regulatory factors can bind to DNA only if a nucleosome is excluded to make the DNA freely accessible, and this creates a boundary for translational positioning. In other cases, regulatory factors can bind to DNA on the surface of the nucleosome, but rotational positioning is important to ensure that the face of DNA with the appropriate contact points is exposed. We discuss the connection between nucleosomal organization and transcription in *Molecular Biology 5.23.4 Nucleosome organization may be changed at the promoter*.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.5.20.11

NUCLEOSOMES

5.20.12 Are transcribed genes organized in nucleosomes?

Key Concepts

- Nucleosomes are found at the same frequency when transcribed genes or nontranscribed genes are digested with micrococcal nuclease.
- Some heavily transcribed genes appear to be exceptional cases that are devoid of nucleosomes.

Attempts to visualize genes during transcription have produced conflicting results. The next two figures show each extreme.

Heavily transcribed chromatin can be seen to be rather extended (too extended to be covered in nucleosomes). In the intensively transcribed genes coding for rRNA, shown in **Figure 20.36**, the extreme packing of RNA polymerases makes it hard to see the DNA. We cannot directly measure the lengths of the rRNA transcripts because the RNA is compacted by proteins, but we know (from the sequence of the rRNA) how long the transcript must be. The length of the transcribed DNA segment, measured by the length of the axis of the "Christmas tree," is ~85% of the length of the rRNA. This means that the DNA is almost completely extended.

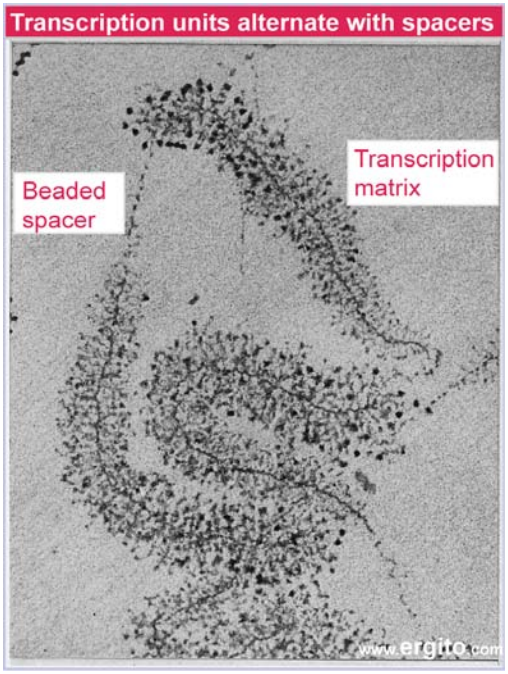

Figure 20.36 The extended axis of an rDNA transcription unit alternates with the only slightly less extended non-transcribed spacer. Photograph kindly provided by Charles Laird.

On the other hand, transcription complexes of SV40 minichromosomes can be extracted from infected cells. They contain the usual complement of histones and display a beaded structure. Chains of RNA can be seen to extend from the minichromosome, as in the example of **Figure 20.37**. This argues that transcription can proceed while the SV40 DNA is organized into nucleosomes. Of course, the SV40 minichromosome is transcribed less intensively than the rRNA genes.

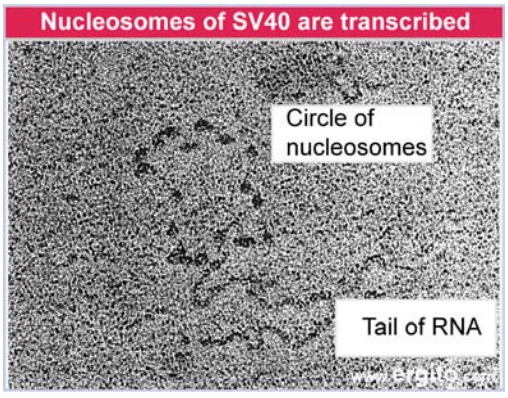

Figure 20.37 An SV40 minichromosome can be transcribed. Photograph kindly provided by Pierre Chambon.

Transcription involves the unwinding of DNA, and may require the fiber to unfold in restricted regions of chromatin. A simple-minded view suggests that some "elbow-room" must be needed for the process. The features of polytene and

lampbrush chromosomes described in *Molecular Biology 5.19 Chromosomes* offer hints that a more expansive structural organization is associated with gene expression.

In thinking about transcription, we must bear in mind the relative sizes of RNA polymerase and the nucleosome. The eukaryotic enzymes are large multisubunit proteins, typically >500 kD. Compare this with the ~260 kD of the nucleosome. **Figure 20.38** illustrates the approach of RNA polymerase to nucleosomal DNA. Even without detailed knowledge of the interaction, it is evident that it involves the approach of two comparable bodies.

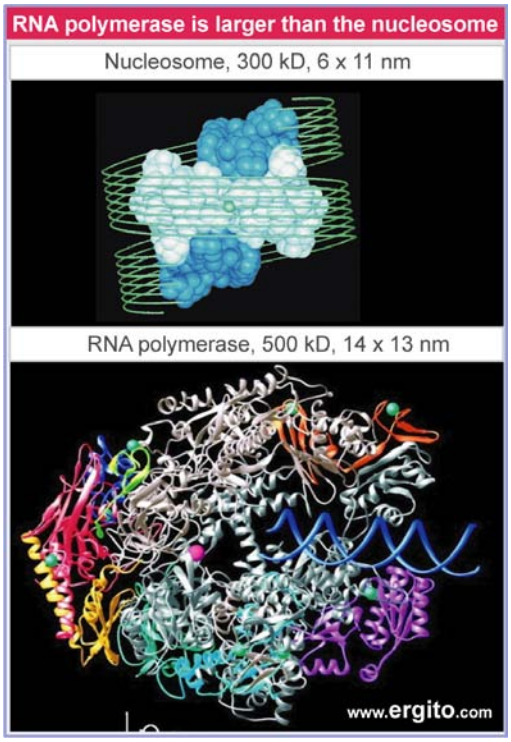

Figure 20.38 RNA polymerase is comparable in size to the nucleosome and might encounter difficulties in following the DNA around the histone octamer.

Consider the two turns that DNA makes around the nucleosome. Would RNA polymerase have sufficient access to DNA if the nucleic acid were confined to this path? During transcription, as RNA polymerase moves along the template, it binds tightly to a region of ~50 bp, including a locally unwound segment of ~12 bp. The need to unwind DNA makes it seem unlikely that the segment engaged by RNA polymerase could remain on the surface of the histone octamer.

It therefore seems inevitable that transcription must involve a structural change. So the first question to ask about the structure of active genes is whether DNA being transcribed remains organized in nucleosomes. If the histone octamers are displaced, do they remain attached in some way to the transcribed DNA (for review see 208)?

One experimental approach is to digest chromatin with micrococcal nuclease, and then to use a probe to some specific gene or genes to determine whether the corresponding fragments are present in the usual 200 bp ladder at the expected concentration. The conclusions that we can draw from these experiments are limited but important. *Genes that are being transcribed contain nucleosomes at the same frequency as nontranscribed sequences.* So genes do not necessarily enter an alternative form of organization in order to be transcribed.

But since the average transcribed gene probably only has a single RNA polymerase at any given moment, this does not reveal what is happening at sites actually engaged by the enzyme. Perhaps they retain their nucleosomes; more likely the nucleosomes are temporarily displaced as RNA polymerase passes through, but reform immediately afterward.

Reviews

208. Kornberg, R. D. and Lorch, Y. (1992). *Chromatin structure and transcription.* Annu. Rev. Cell Biol. 8, 563-587.

NUCLEOSOMES

5.20.13 Histone octamers are displaced by transcription

Key Concepts

- RNA polymerase displaces histone octamers during transcription in a model system, but octamers reassociate with DNA as soon as the polymerase has passed.
- Nucleosomes are reorganized when transcription passes through a gene.

Experiments to test whether an RNA polymerase can transcribe directly through a nucleosome suggest that the histone octamer is displaced by the act of transcription. **Figure 20.39** shows what happens when the phage T7 RNA polymerase transcribes a short piece of DNA containing a single octamer core *in vitro*. The core remains associated with the DNA, but is found in a different location. The core is most likely to rebind to the same DNA molecule from which it was displaced (608).

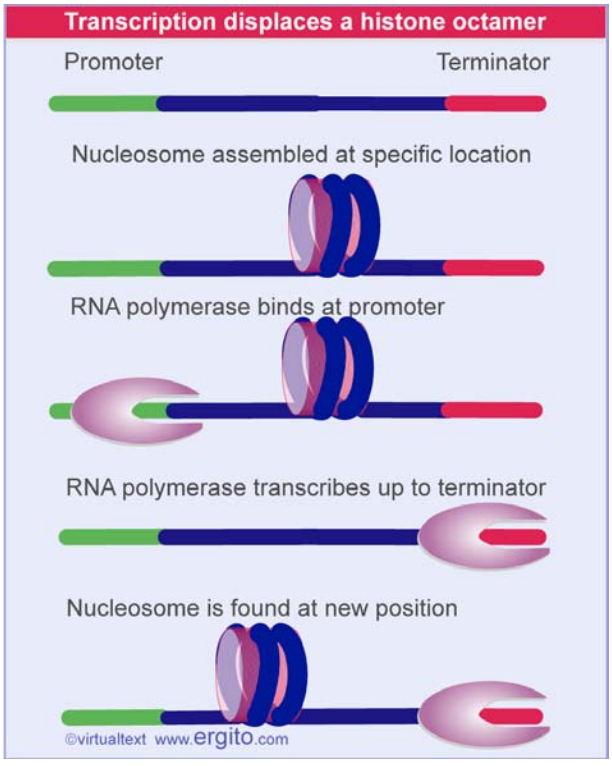

Figure 20.39 A protocol to test the effect of transcription on nucleosomes shows that the histone octamer is displaced from DNA and rebinds at a new position.

Figure 20.40 shows a model for polymerase progression. DNA is displaced as the polymerase enters the nucleosome, but the polymerase reaches a point at which the

DNA loops back and reattaches, forming a closed region. As polymerase advances further, unwinding the DNA, it creates positive supercoils in this loop; the effect could be dramatic, because the closed loop is only ~80 bp, so each base pair through which the polymerase advances makes a significant addition to the supercoiling. In fact, the polymerase progresses easily for the first 30 bp into the nucleosome. Then it proceeds more slowly, as though encountering increasing difficulty in progressing. Pauses occur every 10 bp, suggesting that the structure of the loop imposes a constraint related to rotation around each turn of DNA. When the polymerase reaches the midpoint of the nucleosome (the next bases to be added are essentially at the axis of dyad symmetry), pausing ceases, and the polymerase advances rapidly. This suggests that the midpoint of the nucleosome marks the point at which the octamer is displaced (possibly because positive supercoiling has reached some critical level that expels the octamer from DNA). This releases tension ahead of the polymerase and allows it to proceed. The octamer then binds to the DNA behind the polymerase and no longer presents an obstacle to progress. Probably the octamer changes position without ever completely losing contact with the DNA.

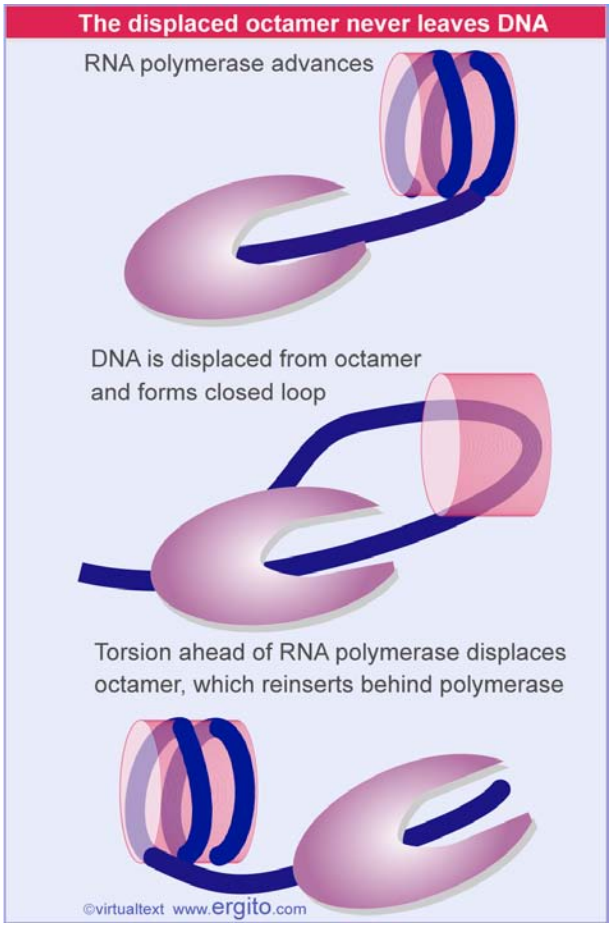

Figure 20.40 RNA polymerase displaces DNA from the histone octamer as it advances. The DNA loops back and attaches (to polymerase or to the octamer) to form a closed loop. As the polymerase proceeds, it generates positive supercoiling ahead. This displaces the octamer, which keeps contact with DNA and/or polymerase, and is inserted behind the RNA polymerase.
This is a static version of an interactive figure; see http://www.ergito.com/main.jsp?bcs=MBIO.5.20.13 to view properly.

Is the octamer released as an intact unit? Crosslinking the proteins of the octamer does not create an obstacle to transcription. Transcription can continue even when crosslinking is extensive enough to ensure that the central regions of the core histones have been linked. This implies that transcription does not require dissociation of the octamer into its component histones, nor is it likely to require any major unfolding of the central structure. However, addition of histone H1 to this system causes a rapid decline in transcription. This suggests two conclusions: the histone octamer (whether remaining present or displaced) functions as an intact unit; and it may be necessary to remove H1 from active chromatin or to modify its interactions in some way.

So a small RNA polymerase can displace a single nucleosome, which reforms behind it, during transcription. Of course, the situation is more complex in a eukaryotic

nucleus. RNA polymerase is very much larger, and the impediment to progress is a string of connected nucleosomes. Overcoming this obstacle requires additional factors that act on chromatin (see *Molecular Biology 5.23 Controlling chromatin structure*).

The organization of nucleosomes may be changed by transcription. **Figure 20.41** shows what happens to the yeast *URA3* gene when it transcribed under control of an inducible promoter (606). Positioning is examined by using micrococcal nuclease to examine cleavage sites relative to a restriction site at the 5′ end of the gene. Initially the gene displays a pattern of nucleosomes that are organized from the promoter for a significant distance across the gene; positioning is lost in the 3′ regions. When the gene is expressed, a general smear replaces the positioned pattern of nucleosomes. So, nucleosomes are present at the same density but are no longer organized in phase. This suggests that transcription destroys the nucleosomal positioning. When repression is reestablished, positioning appears within 10 min (although it is not complete). This result makes the interesting point that the positions of the nucleosomes can be adjusted without replication.

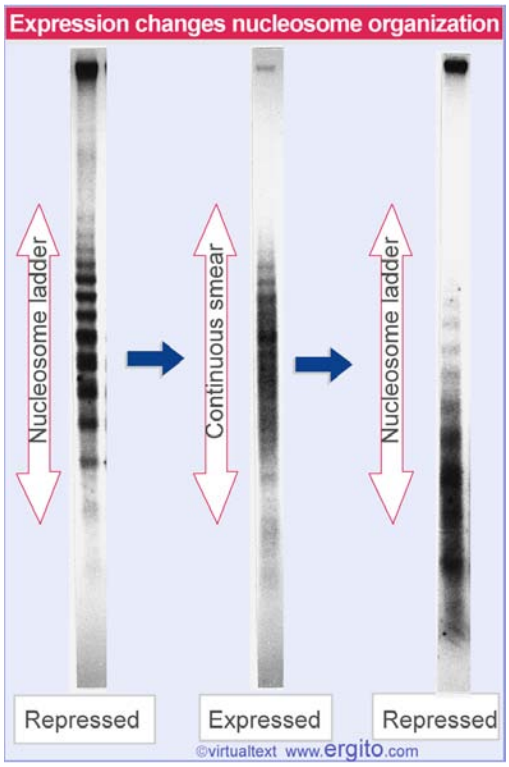

Figure 20.41 The *URA3* gene has positioned nucleosomes before transcription. When transcription is induced, nucleosome positions are randomized. When transcription is repressed, the nucleosomes resume their particular positions. Photograph kindly provided by Fritz Thoma.

The unifying model is to suppose that RNA polymerase displaces histone octamers as it progresses. If the DNA behind the polymerase is available, the octamer reattaches there (possibly or probably never having ever totally lost contact with the DNA. It remains a puzzle how an octamer could retain contact with DNA, without

unfolding or losing components, as an object of even larger size than itself proceeds along the DNA. Perhaps the octamer is "passed back" by making contacts with RNA polymerase). If the DNA is not available, for example, because another polymerase continues immediately behind the first, then the octamer may be permanently displaced, and the DNA may persist in an extended form.

References

606. Cavalli, G. and Thoma, F. (1993). *Chromatin transitions during activation and repression of galactose-regulated genes in yeast*. EMBO J. 12, 4603-4613.

608. Studitsky, V. M., Clark, D. J., and Felsenfeld, G. (1994). *A histone octamer can step around a transcribing polymerase without leaving the template*. Cell 76, 371-382.

NUCLEOSOMES

5.20.14 Nucleosome displacement and reassembly require special factors

Key Concepts

- Ancillary factors are required both for RNA polymerase to displace octamers during transcription and for the histones to reassemble into nucleosomes after transcription.

Displacing nucleosomes from DNA is a key requirement for all stages of transcription. The process has been characterized best at initiation. Active promoters are marked by sites that are hypersensitive to DNAase, because histone octamers have been displaced from DNA (see *Molecular Biology 5.20.15 DNAase hypersensitive sites change chromatin structure*). The removal of the octamers requires remodeling complexes that are recruited by transcription factors and which use energy generated by hydrolysis of ATP to change chromatin structure (see *Molecular Biology 5.23.4 Nucleosome organization may be changed at the promoter*). This means that RNA polymerase starts DNA synthesis on a short stretch of DNA unimpeded by nucleosomes. For it to continue advancing during elongation, the histone octamers ahead of it must be displaced. And then, to avoid leaving naked DNA behind it, the octamers must reform following transcription.

Transcription *in vitro* by RNA polymerase II requires a protein called FACT that behaves like a transcription elongation factor (it is not part of RNA polymerase, but associates with it specifically during the elongation phase of transcription. FACT (an abbreviation for Facilitates Chromatin Transcription) consists of two subunits that are well conserved in all eukaryotes. It is associated with the chromatin of active genes (4108).

When FACT is added to isolated nucleosomes, it causes them to lose H2A·H2B dimers. During transcription *in vitro*, it converts nucleosomes to "hexasomes" that have lost H2A·H2B dimers (4107). This suggests that FACT is part of a mechanism for displacing octamers during transcription. FACT may also be involved in the reassembly of nucleosomes after transcription, because it assists formation of nucleosomes from core histones.

This suggests the model shown in **Figure 20.42**, in which FACT detaches H2A·H2B from a nucleosome in front of RNA polymerase, and then helps to add it to a nucleosome that is reassembling behind the enzyme. Other factors must be required to complete the process. FACT is also required for other reactions in which nucleosomes may be displaced, including DNA replication and repair.

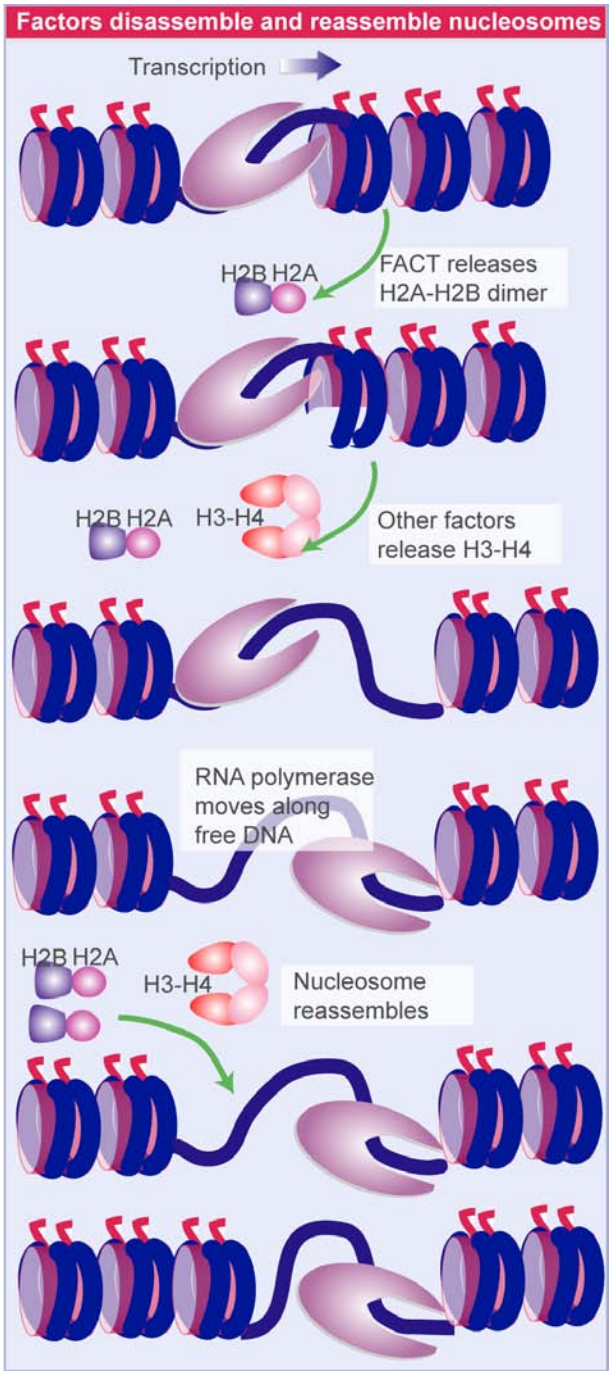

Figure 20.42 Histone octamers are disassembled ahead of transcription to remove nucleosomes. They reform following transcription. Release of H2A.H2B dimers probably initiates the disassembly process.

Other factors are required to maintain the integrity of chromatin in regions that are being transcribed, probably because they are also involved in the disassembly and reassembly of nucleosomes, but we do not yet have detailed information about their functions.

References

4107. Belotserkovskaya, R., Oh, S., Bondarenko, V. A., Orphanides, G., Studitsky, V. M., and Reinberg, D. (2003). *FACT facilitates transcription-dependent nucleosome alteration.* Science 301, 1090-1093.

4108. Saunders, A., Werner, J., Andrulis, E. D., Nakayama, T., Hirose, S., Reinberg, D., and Lis, J. T. (2003). *Tracking FACT and the RNA polymerase II elongation complex through chromatin in vivo.* Science 301, 1094-1096.

NUCLEOSOMES

5.20.15 DNAase hypersensitive sites change chromatin structure

Key Terms

A **hypersensitive site** is a short region of chromatin detected by its extreme sensitivity to cleavage by DNAase I and other nucleases; it comprises an area from which nucleosomes are excluded.

Key Concepts

- Hypersensitive sites are found at the promoters of expressed genes.
- They are generated by the binding of transcription factors that displace histone octamers.

In addition to the general changes that occur in active or potentially active regions, structural changes occur at specific sites associated with initiation of transcription or with certain structural features in DNA. These changes were first detected by the effects of digestion with very low concentrations of the enzyme DNAase I.

When chromatin is digested with DNAase I, the first effect is the introduction of breaks in the duplex at specific, **hypersensitive sites**. Since susceptibility to DNAase I reflects the availability of DNA in chromatin, we take these sites to represent chromatin regions in which the DNA is particularly exposed because it is not organized in the usual nucleosomal structure. A typical hypersensitive site is 100× more sensitive to enzyme attack than bulk chromatin. These sites are also hypersensitive to other nucleases and to chemical agents (for review see 207).

Hypersensitive sites are created by the (tissue-specific) structure of chromatin. Their locations can be determined by the technique of indirect end labeling that we introduced earlier in the context of nucleosome positioning. This application of the technique is recapitulated in **Figure 20.43**. In this case, cleavage at the hypersensitive site by DNAase I is used to generate one end of the fragment, and its distance is measured from the other end that is generated by cleavage with a restriction enzyme.

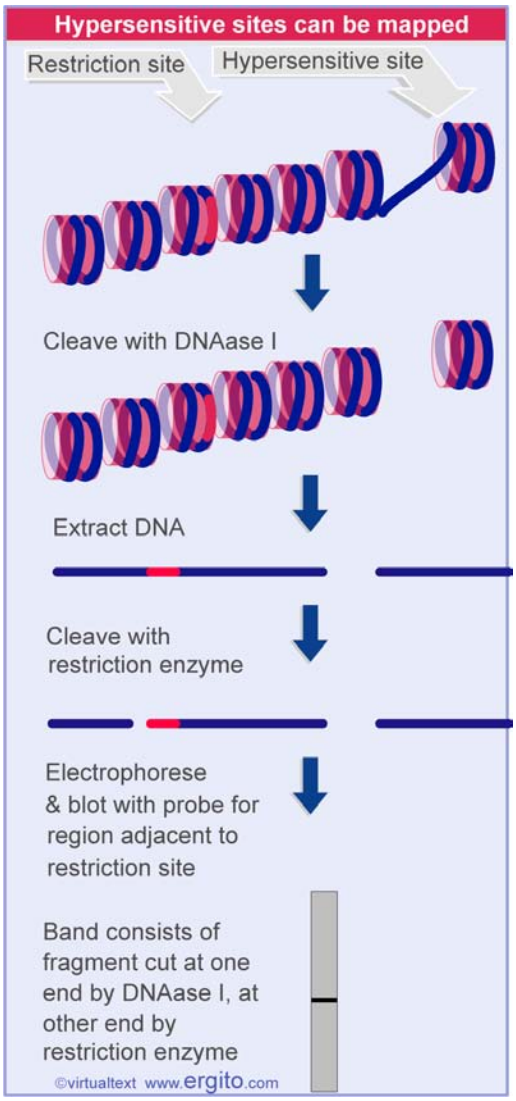

Figure 20.43 Indirect end-labeling identifies the distance of a DNAase hypersensitive site from a restriction cleavage site. The existence of a particular cutting site for DNAase I generates a discrete fragment, whose size indicates the distance of the DNAase I hypersensitive site from the restriction site.

Many of the hypersensitive sites are related to gene expression. Every active gene has a site, or sometimes more than one site, in the region of the promoter. *Most hypersensitive sites are found only in chromatin of cells in which the associated gene is being expressed;* they do not occur when the gene is inactive. The 5′ hypersensitive site(s) appear before transcription begins; and the DNA sequences contained within the hypersensitive sites are required for gene expression, as seen by mutational analysis.

A particularly well-characterized nuclease-sensitive region lies on the SV40 minichromosome. A short segment near the origin of replication, just upstream of the promoter for the late transcription unit, is cleaved preferentially by DNAase I,

micrococcal nuclease, and other nucleases (including restriction enzymes) (610; 611).

The state of the SV40 minichromosome can be visualized by electron microscopy. In up to 20% of the samples, a "gap" is visible in the nucleosomal organization, as evident in **Figure 20.44**. The gap is a region of ~120 nm in length (about 350 bp), surrounded on either side by nucleosomes. The visible gap corresponds with the nuclease-sensitive region. This shows directly that increased sensitivity to nucleases is associated with the exclusion of nucleosomes (3202).

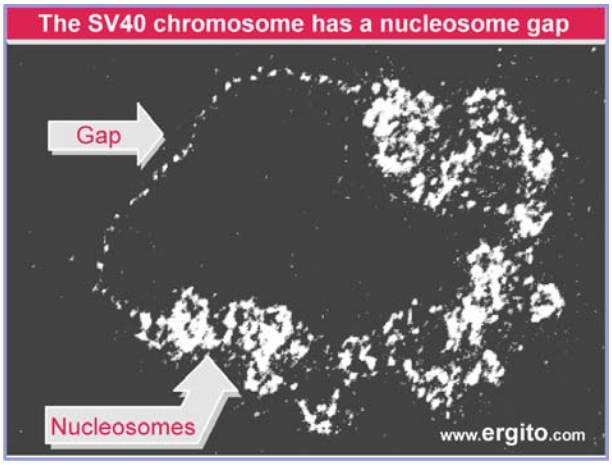

Figure 20.44 The SV40 minichromosome has a nucleosome gap. Photograph kindly provided by Moshe Yaniv.

A hypersensitive site is not necessarily uniformly sensitive to nucleases. **Figure 20.45** shows the maps of two hypersensitive sites.

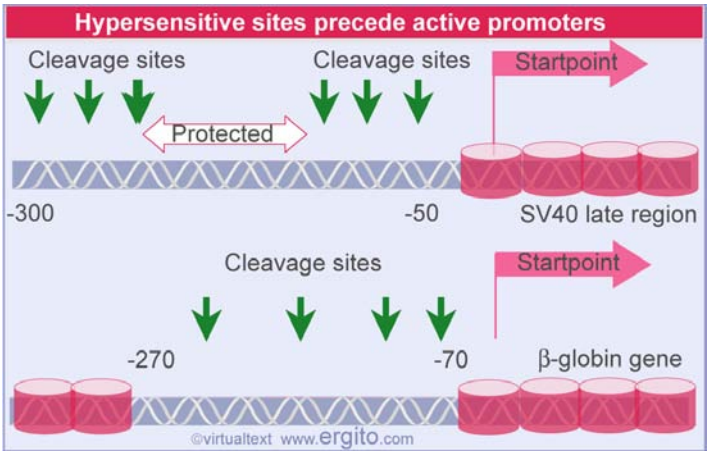

Figure 20.45 The SV40 gap includes hypersensitive sites, sensitive regions, and a protected region of DNA. The hypersensitive site of a chicken β -globin gene comprises a region that is susceptible to several nucleases.

Within the SV40 gap of ~300 bp, there are two hypersensitive DNAase I sites and a "protected" region. The protected region presumably reflects the association of

(nonhistone) protein(s) with the DNA. The gap is associated with the DNA sequence elements that are necessary for promoter function.

The hypersensitive site at the β-globin promoter is preferentially digested by several enzymes, including DNAase I, DNAase II, and micrococcal nuclease. The enzymes have preferred cleavage sites that lie at slightly different points in the same general region. So a region extending from about –70 to –270 is preferentially accessible to nucleases when the gene is transcribable.

What is the structure of the hypersensitive site? Its preferential accessibility to nucleases indicates that it is not protected by histone octamers, but this does not necessarily imply that it is free of protein. A region of free DNA might be vulnerable to damage; and in any case, how would it be able to exclude nucleosomes? We assume that the hypersensitive site results from the binding of specific regulatory proteins that exclude nucleosomes. Indeed, the binding of such proteins is probably the basis for the existence of the protected region within the hypersensitive site.

The proteins that generate hypersensitive sites are likely to be regulatory factors of various types, since hypersensitive sites are found associated with promoters, other elements that regulate transcription, origins of replication, centromeres, and sites with other structural significance. In some cases, they are associated with more extensive organization of chromatin structure. A hypersensitive site may provide a boundary for a series of positioned nucleosomes. Hypersensitive sites associated with transcription may be generated by transcription factors when they bind to the promoter as part of the process that makes it accessible to RNA polymerase (see *Molecular Biology 5.23.4 Nucleosome organization may be changed at the promoter*).

The stability of hypersensitive sites is revealed by the properties of chick fibroblasts transformed with temperature-sensitive tumor viruses. These experiments take advantage of an unusual property: although fibroblasts do not belong to the erythroid lineage, transformation of the cells at the normal temperature leads to activation of the globin genes. The activated genes have hypersensitive sites. If transformation is performed at the higher (nonpermissive) temperature, the globin genes are not activated; and hypersensitive sites do not appear. When the globin genes have been activated by transformation at low temperature, they can be inactivated by raising the temperature. But the hypersensitive sites are retained through at least the next 20 cell doublings (614).

This result demonstrates that acquisition of a hypersensitive site is only one of the features necessary to initiate transcription; and it implies that the events involved in establishing a hypersensitive site are distinct from those concerned with perpetuating it. Once the site has been established, it is perpetuated through replication in the absence of the circumstances needed for induction. Could some specific intervention be needed to abolish a hypersensitive site?

Reviews

207. Gross, D. S. and Garrard, W. T. (1988). *Nuclease hypersensitive sites in chromatin.* Annu. Rev. Biochem. 57, 159-197.

References

610. Varshavsky, A. J., Sundin, O., and Bohn, M. J. (1978). *SV40 viral minichromosome: preferential exposure of the origin of replication as probed by restriction endonucleases*. Nucleic Acids Res. 5, 3469-3479.

611. Scott, W. A. and Wigmore, D. J. (1978). *Sites in SV40 chromatin which are preferentially cleaved by endonucleases*. Cell 15, 1511-1518.

614. Groudine, M., and Weintraub, H. (1982). *Propagation of globin DNAase I-hypersensitive sites in absence of factors required for induction: a possible mechanism for determination*. Cell 30, 131-139.

3202. Moyne, G., Harper, F., Saragosti, S., and Yaniv, M. (1982). *Absence of nucleosomes in a histone-containing nucleoprotein complex obtained by dissociation of purified SV40 virions.* Cell 30, 123-130.

NUCLEOSOMES

5.20.16 Domains define regions that contain active genes

Key Terms

A **domain** of a chromosome may refer *either* to a discrete structural entity defined as a region within which supercoiling is independent of other domains; *or* to an extensive region including an expressed gene that has heightened sensitivity to degradation by the enzyme DNAase I.

Key Concepts

- A domain containing a transcribed gene is defined by increased sensitivity to degradation by DNAase I.

A region of the genome that contains an active gene may have an altered structure. The change in structure precedes, and is different from, the disruption of nucleosome structure that may be caused by the actual passage of RNA polymerase.

One indication of the change in structure of transcribed chromatin is provided by its increased susceptibility to degradation by DNAase I. DNAase I sensitivity defines a chromosomal **domain**, a region of altered structure including at least one active transcription unit, and sometimes extending farther. (Note that use of the term "domain" does not imply any necessary connection with the structural domains identified by the loops of chromatin or chromosomes.)

When chromatin is digested with DNAase I, it is eventually degraded into acid-soluble material (very small fragments of DNA). The progress of the overall reaction can be followed in terms of the proportion of DNA that is rendered acid soluble. *When only 10% of the total DNA has become acid soluble, more than 50% of the DNA of an active gene has been lost.* This suggests that active genes are preferentially degraded (613).

The fate of individual genes can be followed by quantitating the amount of DNA that survives to react with a specific probe. The protocol is outlined in **Figure 20.46**. The principle is that the loss of a particular band indicates that the corresponding region of DNA has been degraded by the enzyme.

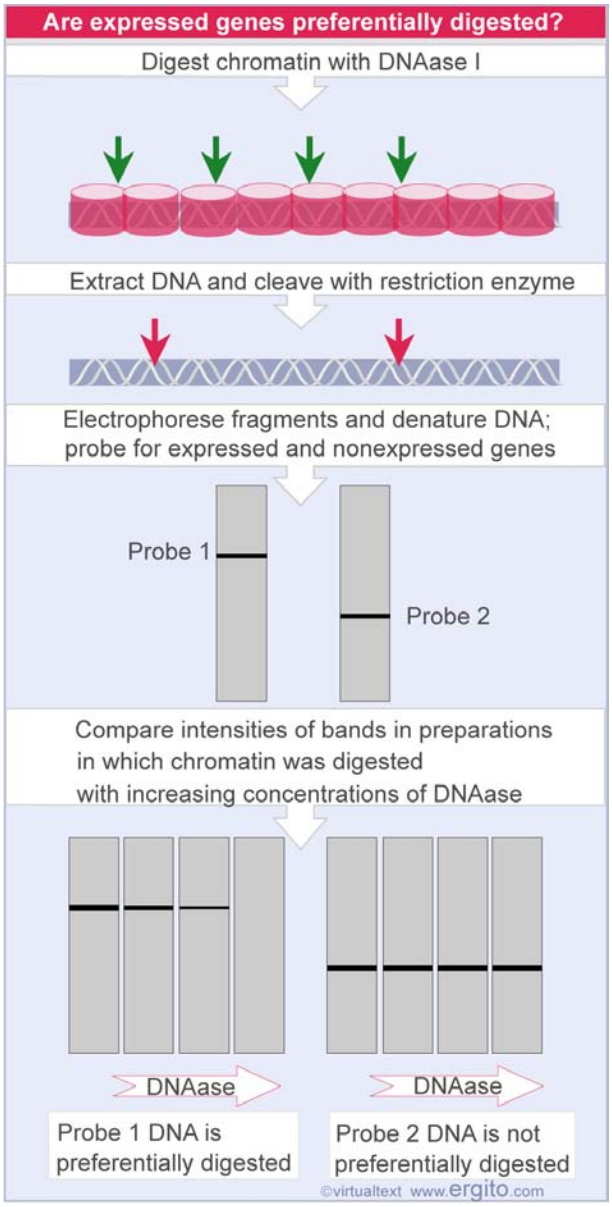

Figure 20.46 Sensitivity to DNAase I can be measured by determining the rate of disappearance of the material hybridizing with a particular probe.

Figure 20.47 shows what happens to β-globin genes and an ovalbumin gene in chromatin extracted from chicken red blood cells (in which globin genes are expressed and the ovalbumin gene is inactive). The restriction fragments representing the β-globin genes are rapidly lost, while those representing the ovalbumin gene show little degradation. (The ovalbumin gene in fact is digested at the same rate as the bulk of DNA.)

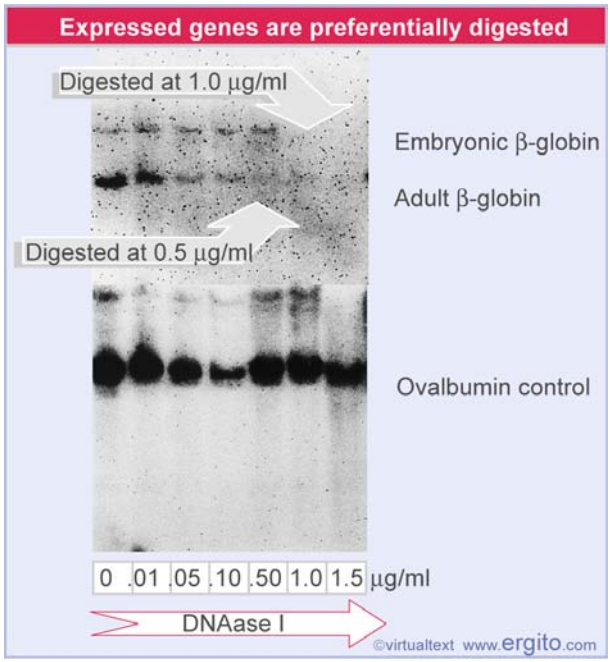

Figure 20.47 In adult erythroid cells, the adult β -globin gene is highly sensitive to DNAase I digestion, the embryonic β -globin gene is partially sensitive (probably due to spreading effects), but ovalbumin is not sensitive. Data kindly provided by Harold Weintraub.

So the bulk of chromatin is relatively resistant to DNAase I and contains nonexpressed genes (as well as other sequences). *A gene becomes relatively susceptible to the enzyme specifically in the tissue(s) in which it is expressed.*

Is preferential susceptibility a characteristic only of rather actively expressed genes, such as globin, or of all active genes? Experiments using probes representing the entire cellular mRNA population suggest that all active genes, whether coding for abundant or for rare mRNAs, are preferentially susceptible to DNAase I. (However, there are variations in the degree of susceptibility.) Since the rarely expressed genes are likely to have very few RNA polymerase molecules actually engaged in transcription at any moment, this implies that the sensitivity to DNAase I does not result from the act of transcription, but is a feature of *genes that are able to be transcribed.*

What is the extent of the preferentially sensitive region? This can be determined by using a series of probes representing the flanking regions as well as the transcription unit itself. The sensitive region always extends over the entire transcribed region; an additional region of several kb on either side may show an intermediate level of sensitivity (probably as the result of spreading effects).

The critical concept implicit in the description of the domain is that a region of high sensitivity to DNAase I extends over a considerable distance. Often we think of regulation as residing in events that occur at a discrete site in DNA – for example, in the ability to initiate transcription at the promoter. Even if this is true, such regulation must determine, or must be accompanied by, a more wide-ranging change in

structure. This is a difference between eukaryotes and prokaryotes.

References

613. Stalder, J. et al. (1980). *Tissue-specific DNA cleavage in the globin chromatin domain introduced by DNAase I*. Cell 20, 451-460.

NUCLEOSOMES

5.20.17 An LCR may control a domain

Key Terms

The **locus control region (LCR)** that is required for the expression of several genes in a domain.

A **domain** of a chromosome may refer *either* to a discrete structural entity defined as a region within which supercoiling is independent of other domains; *or* to an extensive region including an expressed gene that has heightened sensitivity to degradation by the enzyme DNAase I.

Key Concepts

- An LCR is located at the 5′ end of the domain and consists of several hypersensitive sites.

Every gene is controlled by its promoter, and some genes also respond to enhancers (containing similar control elements but located farther away) as discussed in *Molecular Biology 5.21 Promoters and enhancers*. However, these local controls are not sufficient for all genes. In some cases, a gene lies within a domain of several genes all of which are influenced by regulatory elements that act on the whole domain. The existence of these elements was identified by the inability of a region of DNA including a gene and all its known regulatory elements to be properly expressed when introduced into an animal as a transgene.

The best characterized example of a regulated gene cluster is provided by the mouse β-globin genes. Recall from **Figure 4.3** that the α-globin and β-globin genes in mammals each exist as clusters of related genes, expressed at different times during embryonic and adult development. These genes are provided with a large number of regulatory elements, which have been analyzed in detail. In the case of the adult human β-globin gene, regulatory sequences are located both 5′ and 3′ to the gene and include both positive and negative elements in the promoter region, and additional positive elements within and downstream of the gene.

But a human β-globin gene containing all of these control regions is never expressed in a transgenic mouse within an order of magnitude of wild-type levels. Some further regulatory sequence is required. Regions that provide the additional regulatory function are identified by DNAase I hypersensitive sites that are found at the ends of the cluster. The map of **Figure 20.48** shows that the 20 kb upstream of the ε-gene contains a group of 5 sites; and there is a single site 30 kb downstream of the β-gene. Transfecting various constructs into mouse erythroleukemia cells shows that sequences between the individual hypersensitive sites in the 5′ region can be removed without much effect, but that removal of any of the sites reduces the overall level of expression.

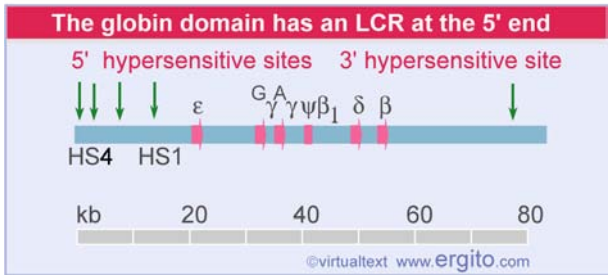

Figure 20.48 A globin domain is marked by hypersensitive sites at either end. The group of sites at the 5′ side constitutes the LCR and is essential for the function of all genes in the cluster.

The 5′ regulatory sites are the primary regulators, and the cluster of hypersensitive sites is called the **LCR** (locus control region) (for review see 3182). We do not know whether the 3′ site has any function. The LCR is absolutely required for expression of each of the globin genes in the cluster. Each gene is then further regulated by its own specific controls. Some of these controls are autonomous: expression of the ε- and γ-genes appears intrinsic to those loci in conjunction with the LCR. Other controls appear to rely upon position in the cluster, which provides a suggestion that *gene order* in a cluster is important for regulation (615).

The entire region containing the globin genes, and extending well beyond them, constitutes a chromosomal **domain**. It shows increased sensitivity to digestion by DNAase I (see **Figure 20.46**). Deletion of the 5′ LCR restores normal resistance to DNAase over the whole region. Two models for how an LCR works propose that its action is required in order to activate the promoter, or alternatively, to increase the rate of transcription from the promoter (3183; 3181). The exact nature of the interactions between the LCR and the individual promoters has not yet been fully defined.

Does this model apply to other gene clusters? The α-globin locus has a similar organization of genes that are expressed at different times, with a group of hypersensitive sites at one end of the cluster, and increased sensitivity to DNAase I throughout the region. Only a small number of other cases are known in which an LCR controls a group of genes.

Reviews

3181. Bulger, M. and Groudine, M. (1999). *Looping versus linking: toward a model for long-distance gene activation.* Genes Dev. 13, 2465-2477.

3182. Grosveld, F., Antoniou, M., Berry, M., De Boer, E., Dillon, N., Ellis, J., Fraser, P., Hanscombe, O., Hurst, J., and Imam, A. (1993). *The regulation of human globin gene switching.* Philos. Trans. R. Soc. Lond. B Biol. Sci. 339, 183-191.

References

615. van Assendelft, G. B., Hanscombe, O., Grosveld, F., and Greaves, D. R. (1989). *The β -globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner*. Cell 56, 969-977.

3183. Gribnau, J., de Boer, E., Trimborn, T., Wijgerde, M., Milot, E., Grosveld, F., and Fraser, P. (1998). *Chromatin interaction mechanism of transcriptional control in vitro.* EMBO J. 17, 6020-6027.

NUCLEOSOMES

5.20.18 Summary

All eukaryotic chromatin consists of nucleosomes. A nucleosome contains a characteristic length of DNA, usually ~200 bp, wrapped around an octamer containing two copies each of histones H2A, H2B, H3, and H4. A single H1 protein is associated with each nucleosome. Virtually all genomic DNA is organized into nucleosomes. Treatment with micrococcal nuclease shows that the DNA packaged into each nucleosome can be divided operationally into two regions. The linker region is digested rapidly by the nuclease; the core region of 146 bp is resistant to digestion. Histones H3 and H4 are the most highly conserved and an $H3_2 \cdot H4_2$ tetramer accounts for the diameter of the particle. The H2A and H2B histones are organized as two H2A·H2B dimers. Octamers are assembled by the successive addition of two H2A·H2B dimers to the $H3_2 \cdot H4_2$ kernel.

The path of DNA around the histone octamer creates –1.65 supercoils. The DNA "enters" and "leaves" the nucleosome in the same vicinity, and could be "sealed" by histone H1. Removal of the core histones releases –1.0 supercoils. The difference can be largely explained by a change in the helical pitch of DNA, from an average of 10.2 bp/turn in nucleosomal form to 10.5 bp/turn when free in solution. There is variation in the structure of DNA from a periodicity of 10.0 bp/turn at the nucleosome ends to 10.7 bp/turn in the center. There are kinks in the path of DNA on the nucleosome.

Nucleosomes are organized into a fiber of 30 nm diameter which has 6 nucleosomes per turn and a packing ratio of 40. Removal of H1 allows this fiber to unfold into a 10 nm fiber that consists of a linear string of nucleosomes. The 30 nm fiber probably consists of the 10 nm fiber wound into a solenoid. The 30 nm fiber is the basic constituent of both euchromatin and heterochromatin; nonhistone proteins are responsible for further organization of the fiber into chromatin or chromosome ultrastructure.

There are two pathways for nucleosome assembly. In the replication-coupled pathway, the PCNA processivity subunit of the replisome recruits CAF-1, which is a nucleosome assembly factor. CAF-1 assists the deposition of $H3_2 \cdot H4_2$ tetramers onto the daughter duplexes resulting from replication. The tetramers may be produced either by disruption of existing nucleosomes by the replication fork or as the result of assembly from newly synthesized histones. Similar sources provide the H2A·H2B dimers that then assemble with the $H3_2 \cdot H4_2$ tetramer to complete the nucleosome. Because the $H3_2 \cdot H4_2$ tetramer and the H2A·H2B dimers assemble at random, the new nucleosomes may include both pre-existing and newly synthesized histones.

RNA polymerase displaces histone octamers during transcription. Nucleosomes reform on DNA after the polymerase has passed, unless transcription is very intensive (such as in rDNA) when they may be displaced completely. The replication-independent pathway for nucleosome assembly is responsible for replacing histone octamers that have been displaced by transcription. It uses the histone variant H3.3 instead of H3. A similar pathway, with another alternative to H3, is used for assembling nucleosomes at centromeric DNA sequences following

replication.

Two types of changes in sensitivity to nucleases are associated with gene activity. Chromatin capable of being transcribed has a generally increased sensitivity to DNAase I, reflecting a change in structure over an extensive region that can be defined as a domain containing active or potentially active genes. Hypersensitive sites in DNA occur at discrete locations, and are identified by greatly increased sensitivity to DNAase I. A hypersensitive site consists of a sequence of ~200 bp from which nucleosomes are excluded by the presence of other proteins. A hypersensitive site forms a boundary that may cause adjacent nucleosomes to be restricted in position. Nucleosome positioning may be important in controlling access of regulatory proteins to DNA.

Hypersensitive sites occur at several types of regulators. Those that regulate transcription include promoters, enhancers, and LCRs. Other sites include origins for replication and centromeres. A promoter or enhancer acts on a single gene, but an LCR contains a group of hypersensitive sites and may regulate a domain containing several genes.