UNRAVELING THE CHROMOSOME

E. Morton Bradbury

Central to biology is an understanding of the organization, structure, and functions of the chromosomes of higher organisms. Chromosomes contain the DNA molecules of the genome and are themselves contained within the cell nuclei of all eukaryotes, from single-celled yeast all the way up the evolutionary ladder to human beings. As pointed out by David Galas (pages 164–165 of "Mapping the Genome"), to understand the functions of the multitude of protein-coding and noncoding DNA sequences that will be determined by the Human Genome Project, we will need detailed knowledge of the three-dimensional structure of chromosomes and the structural changes that chromosomes undergo during the various phases of the cell cycle. Major advances in biology will be at the interfaces between the Human Genome Project, structural biology, and molecular biology of the cell.

The size of the human genome suggests the magnitude of the problem. The diploid human genome contains 6×10^9 base pairs or 204 centimeters of DNA molecules packaged into 46 chromosomes. It is generally believed that each chromosome contains a single DNA molecule several centimeters in length.

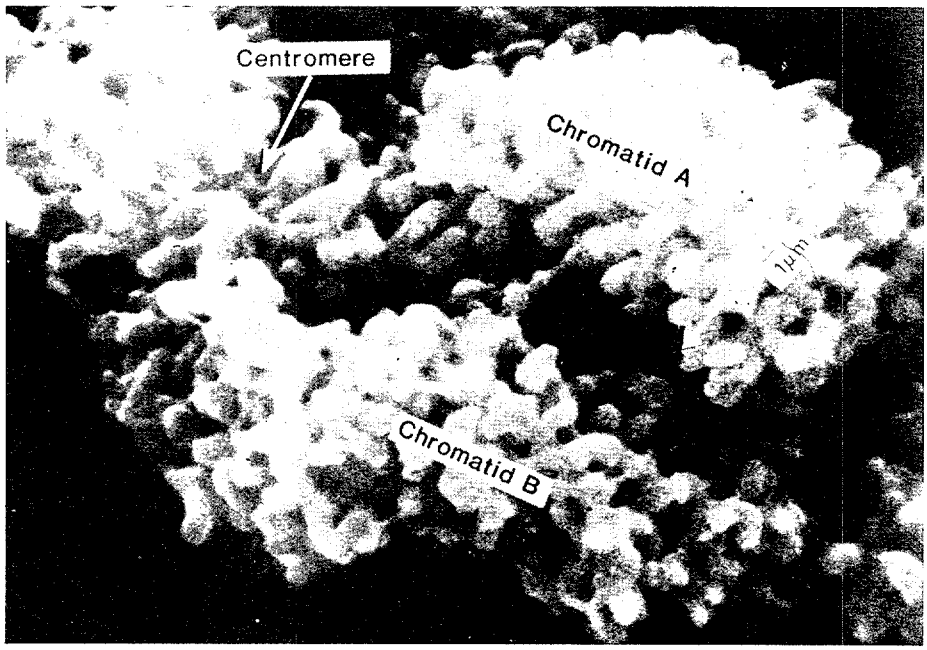


Figure 1. Human Metaphase Chromosome
A scanning transmission electron micrograph of a metaphase chromosome showing two sister chromatids attached at the centromeres. Each compact projection is thought to be a long loop of DNA (see Figure 2) packaged along with various proteins into a thick chromatin fiber. (Reprinted courtesy of U.K. Laemmli, Université de Genève.)

Studies of the yeast *S. cerevisiae*, a lower eukaryote that can be easily manipulated, have revealed three chromosomal elements that are essential to the faithful replication of each chromosome and to the subsequent separation of the two duplicate chromosomes into daughter cells during cell division. These are: (1) the very ends of chromosomes, called the telomeres; (2) a central region of constriction called the centromere that, after replication of a chromosome, is the last point of attachment between the resulting pair of sister chromatids; and (3) a DNA sequence required to initiate DNA replication, called an origin of replication.

Figure 1 is a scanning transmission electron micrograph of a human metaphase chromosome, the highly condensed structure adopted by the chromosome during metaphase, one of the last phases of cell division. The chromosome has already replicated into two sister chromatids. The centromere connecting the sister chromatids (seen in the micrograph as a region of constriction) provides the point of attachment

for the spindle apparatus that contracts and separates the replicated chromosomes into the daughter cells. The telomeres at the ends of each chromatid contain tandem repeated DNA sequences that cap, protect, and maintain the linear DNA ends of the chromosomes during replication.

Each of the 46 human chromosomes can be identified during metaphase by its length, the location of its centromere, and the particular banding pattern produced by staining the DNA of that chromosome. (Banding patterns can be seen in "Chromosomes: The Sites of Hereditary Information" in "Understanding Inheritance.") The origins of the distinctive banding patterns are not well understood but probably reflect a reproducible pattern of DNA folding induced by DNA-protein interactions specific to each chromosome. The DNA molecule is very tightly wound during metaphase. For example, human chromosome 16 is 2.5 micrometers long, whereas the DNA molecule in each sister chromatid is 3.7 centimeters long. In other words, the packing ratio of the linear DNA molecule in the metaphase chromosome is 15,000 to 1.

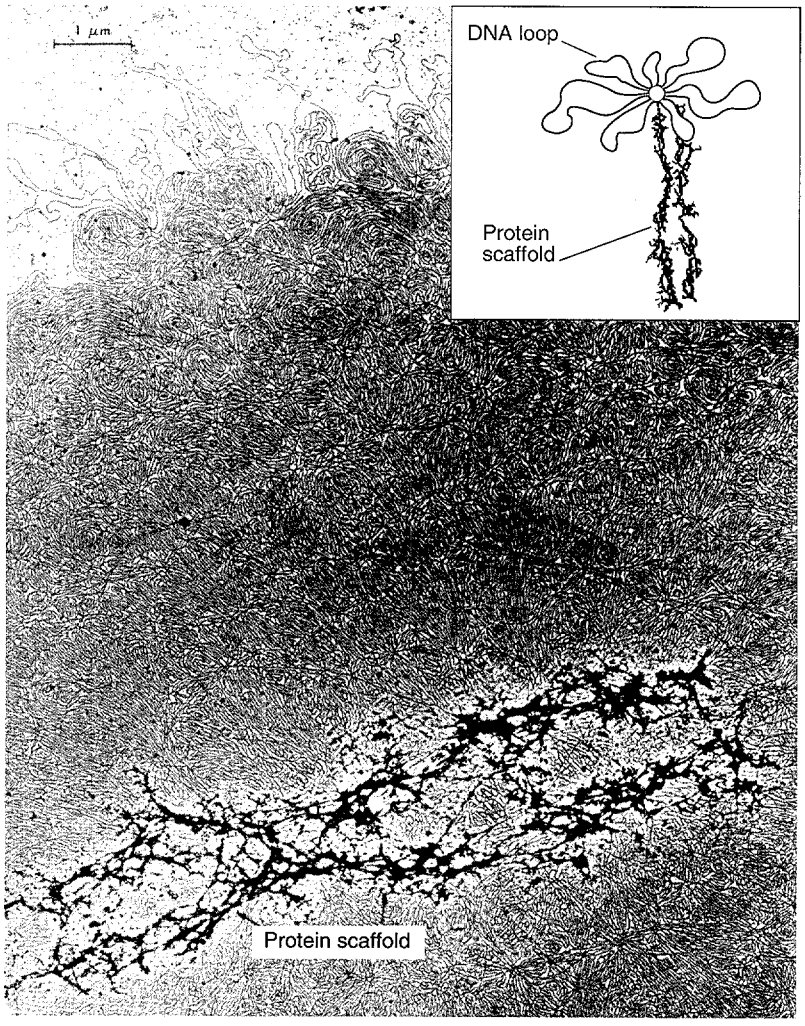


Figure 2. Chromosome Loops and Protein Scaffold

Above is a metaphase chromosome depleted of almost all chromosomal proteins. The remaining 2 to 3 percent of the proteins form a scaffold that retains the shape of the intact chromosome. Around the scaffold is a halo of loops of naked DNA. Each loop appears to begin and end at the same point along the protein scaffold (see insert). The number and sizes of these loops suggest that each may contain a single gene or a group of linked genes. (Reprinted courtesy of U.K. Laemmli, Université de Genève.)

Chromosomal DNA Loops

When chromosomal material is isolated from the nucleus, the long DNA molecules are found to be associated with chromosomal proteins, whose weight is up to twice that of the DNA. The five histones, the many copies of which are equal in weight to that of DNA, are found in all eukaryotes and as explained below are involved in packaging the DNA in the chromosomes. The non-histone proteins are a heterogeneous group and many are associated with the various chromosome functions, such as replication, gene expression, and chromosome organization. Among the latter are a small group that bind most tightly to the DNA and form a scaffold for the chromosome. This protein scaffold has been made visible by gently treating metaphase chromosomes with detergents to remove the histones and most other nonhistone proteins. The remarkable structure that remains is shown in Figure 2. The residual protein scaffold, or "ghost," of the metaphase chromosome is surrounded by a halo of DNA. At higher

resolution DNA loops can be observed to emerge from and return to the same point on the protein scaffold (see inset in Figure 2).

Two major scaffold proteins have been isolated, Sc1 and Sc2. Sc1 has been identified as topoisomerase II, an enzyme that relaxes supercoiled DNA by cutting through both strands of the DNA, thereby enabling the cut DNA ends to rotate, and then resealing the cut. The cuts made by topoisomerase II are essential for the separation of sister chromatids to the daughter cells.

The DNA loops in Figure 2 range in size from 5,000 to 120,000 base pairs and have an average size of about 50,000 base pairs. Thus the haploid human genome of 3×10^9 base pairs of DNA corresponds to 60,000 loops, which is close to the estimated numbers of genes, 50,000 to 100,000, in the human genome. Perhaps each DNA loop contains one or a small number of linked genes and therefore serves as both a genetic and a structural unit of eukaryotic chromosomes. This tantalizing conjecture was first made in 1978, and although it remains unproven, evidence in its favor has been accumulating.

Chromatin Contains a Repeating Subunit Structure

Having looked at some of the largest structural features of the chromosome, we now turn to what we know about the small, repeating substructures within a chromosome. DNA with its associated chromosomal proteins, histones, and nonhistone proteins, is called chromatin. In 1973 chromatin in isolated nuclei was first digested with micrococcal nuclease, an enzyme that cuts double-stranded DNA. The digestion yielded a ladder of DNA lengths in multiples of about 190 to 200 base pairs. Evidently DNA sequences spaced by 190 to 200 base pairs were more accessible to attack by micrococcal nuclease than the intervening DNA. This seminal observation showed that chromatin contained a simple, repeating subunit, known as the nucleosome.

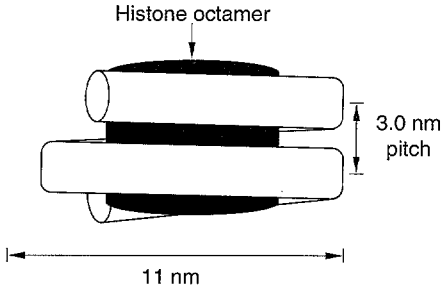


Figure 3. Nucleosome Core Particle
Structure of the nucleosome core particle determined from neutron scattering. The core particle is a flat disc, 100 angstroms in diameter and 55 to 60 angstroms thick.

For most somatic tissues, the nucleosome contains three elements, a stretch of DNA containing 195±5 base pairs, one copy of the histone octamer $[(H3_2H4_2)(H2A,H2B)_2]$ and one copy of the histone H1. More prolonged micrococcal nuclease digestion reduces the length of the DNA in the nucleosome, thereby creating a slightly smaller unit, called the chromatosome, which contains 168±2 base pairs of DNA, the histone octamer, and H1. Such digestion often reduces the nucleosome to an even smaller unit contained within the chromatosome and called the nucleosome core particle. It contains 146±1 base pairs of DNA and the histone octamer (see Figure 3).

The nucleosome core particle has been obtained in large quantities and subjected to extensive structural studies. In 1974 neutron-scattering studies of the core particle in aqueous solution showed that it was a flat disc of diameter 100 angstroms and thickness 55 to 60 angstroms, with 1.7 turns of DNA coiled on the outside of a core of the histone octamer at a pitch of about 30 angstroms

(Figure 3). Subsequent x-ray-diffraction studies of crystallized core particles achieved a resolution of 6 to 7 angstroms. The crystal structure (Figure 4) not only confirmed the lower resolution solution structure achieved by neutron scattering but also showed that histones are in contact with the minor groove of DNA and leave the major groove available for interactions with the proteins that regulate gene expression and other DNA functions. The 7-angstrom-resolution crystal structure also revealed that DNA does not bend uniformly but rather bends gently and then more sharply around the histone octamer. Such a path implies that flexibility, or bendability, of DNA may be sequence-dependent and that the underlying DNA sequence along the molecule may determine the positions of some nucleosomes. The most recent work on nucleosome positioning shows that the bulk of nucleosome core particles are able to move along the DNA molecule between a cluster of positions separated by about 10 base pairs. This mobility is probably required during DNA replication and transcription to allow DNA polymerases and other enzymes access to specific DNA sequences.

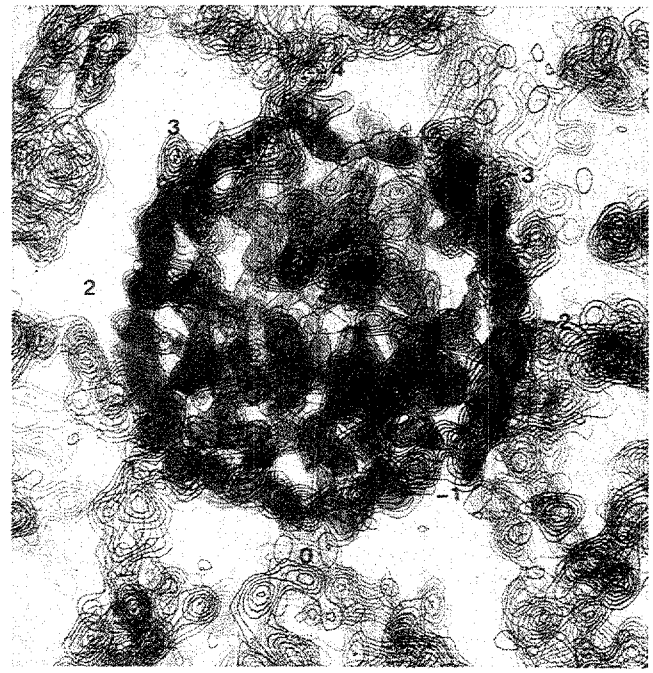


Figure 4. Crystal Structure of Core Particle

The structure of the nucleosome core particle as determined by x-ray diffraction is shown above. At a resolution of 6 to 7 angstroms, this top view of the core particle shows that the DNA (brown) does not follow a smooth path around the histone octamer (blue and turquoise) but rather bends sharply and then more gently. (Reprinted courtesy of Uberbacher and Bunick, Oak Ridge National Laboratory.)

Despite considerable effort to achieve higher resolution, the best data for the core particle structure is at a resolution of about 6 angstroms. However, the crystal structure of the isolated histone octamer has been solved to the higher resolution of 3.3 angstroms. This structure shows shapes of the individual histones and the nature of interhistone interaction of most but not all of the histone polypeptide chains. In particular, the basic N-terminal domains, comprising some 20 to 25 percent of the histone octamer, are not "seen" in the crystal structure, probably because they bind to DNA, and in the absence of DNA, they are disordered. These N-terminal domains contain all of the sites of the cell-cycle-dependent acetylation of lysines and phosphorylation of serines or threonines. Acetylation of lysine converts it from a positively charged residue, which can therefore bind to DNA, to a neutral acetylysine. It has been shown first that lysine acetylation is strictly correlated with transcription and DNA replication, and more recently, that histone acetylation drives the uncoiling of part of the DNA from the nucleosome to allow the initiation and progression of DNA replication and transcription.

Chromatosomes and Nucleosomes

A model of the structure of the chromatosome (Figure 5) has been inferred from the structures of the nucleosome core particle and the histone H1. The core particle has 1.7 turns of DNA at a pitch of 3.0 nanometers (30 angstroms) coiled around the histone octamer. Consequently, the chromatosome's 168 base pairs of DNA are long

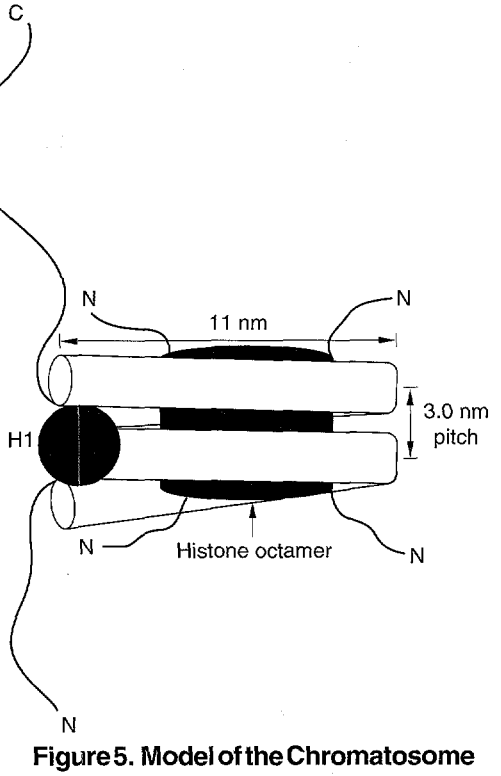


Figure 5. Model of the Chromatosome
The model includes the nucleosome core particle, an extra stretch of DNA, and the histone H1. The DNA makes two complete turns around the histone octamer, and H1 is bound to the outside of the coil at the place where the coil begins and ends. In this position H1 might server to modulate long-range interactions that modify chromosome structure during the cell cycle.

enough to complete two turns of DNA around the histone octamer. The chromatosome also includes the fifth histone H1. In the model structure shown in Figure 5, the histone H1 is bound to the outside of the coiled DNA where it might serve to modulate long-range interactions associated with reversible changes in chromosome structure during the cell cycle. During cell division chromosomes become more and more condensed until they reach metaphase. Then, when cell division is completed and the daughter cells enter interphase, the chromosomes assume a less-condensed configuration (see "Mitosis" in "Understanding Inheritance"). The long, flexible "arms" of H1 undergo a pattern of phosphorylations through this cycle, which may well modulate the long-range interactions required to coordinate these structural changes in the chromosomes. In support of this hypothesis is the fact that an increase in H1 phosphorylation has been correlated with the process of chromosome condensation to metaphase chromosomes. To describe the nucleosome beyond the model for the chromatosome requires a knowledge of the paths of the DNA that link one nucleosome to another. Our present lack of knowledge about those paths impedes our ability to pin down the higher-order chromatin structures that make up the chromosome.

Higher-Order Chromatin Structures

Although higher-order structures of chromatin cannot be resolved in the chromosome itself, they can be studied in solution. Chromatin, when placed in low ionic strength, 10-millimolar NaCl, forms a 10-nanometer-diameter fibril of nucleosomes, which is sometimes referred to as "beads on a string." This form is also observed when chromatin spills out of lysed nuclei. Neutron-scattering studies of the 10-nanometer chromatin fibril give a mass per unit length equivalent to one nucleosome per 10±2 nanometers of fibril, or a DNA packing ratio of between 6 and 7 to 1. When ionic strength is increased to 150-millimolar NaCl, corresponding to normal physiological conditions, the 10-nanometer fibril undergoes a transition to the "30-nanometer" fibril. Neutron-scattering studies indicate that the diameter for this fibril in solution is 34 nanometers and the mass-per-unit length is equivalent to 6 to 7 nucleosomes per 11 nanometers of fibril, or a DNA packing ratio of between 40 and 50 to 1. Figure 6 shows the simplest model of the 34-nanometer fibril that is consistent with available structural data: it is a supercoil or solenoid of 6 to 7 radially arranged disc-shaped nucleosomes with a pitch of 11.0 nanometers and a diameter of 34 nanometers. Basic questions concerning the location of histone H1 and the linker DNA connecting the nucleosomes remain unanswered.

Packaging of Chromosome Loops

With these higher order chromatin structures in mind, we can imagine how the large transverse DNA loops present in the histone-depleted metaphase chromosome (see Figure 2) might be packaged in the normal chromosome. Since the average size of the

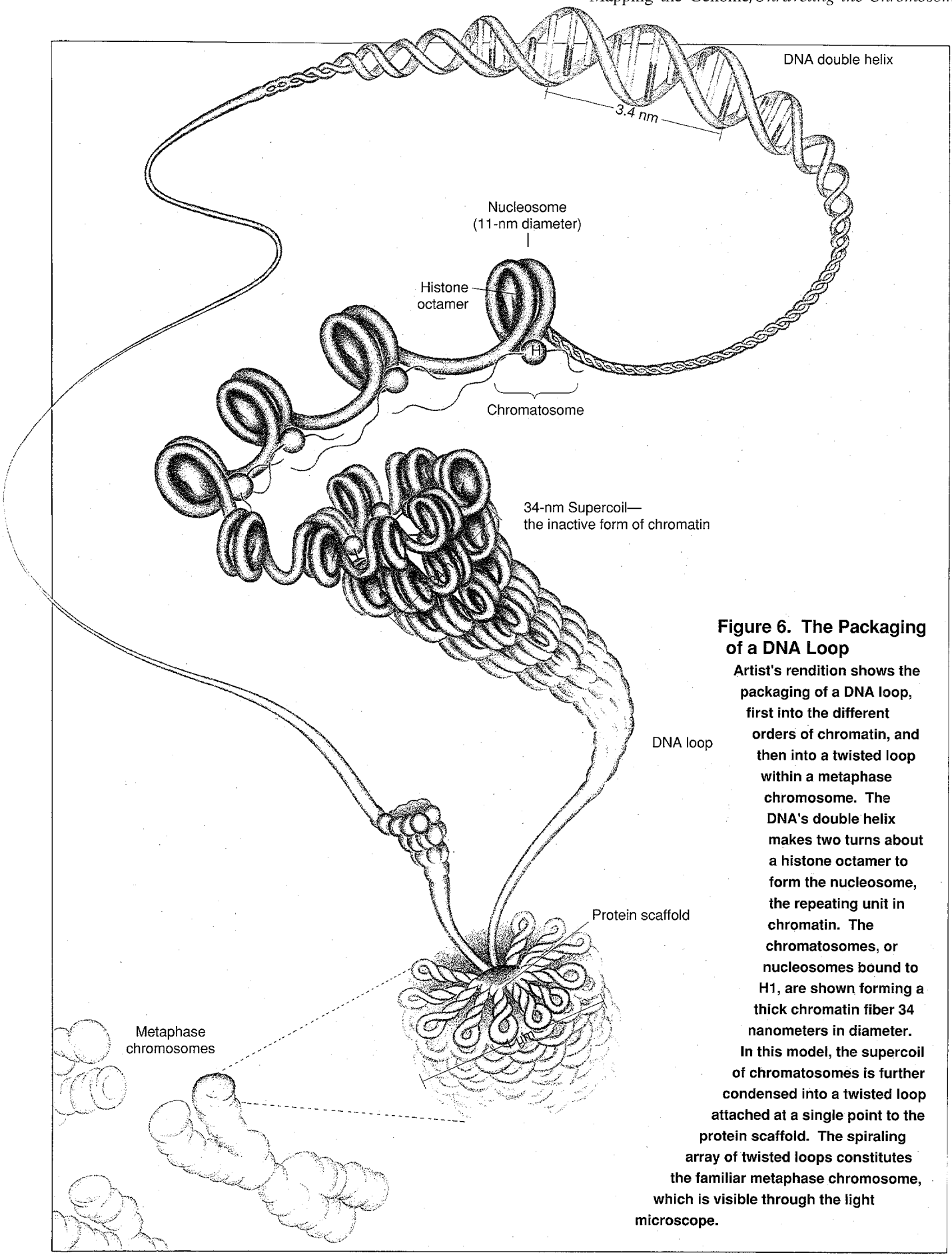


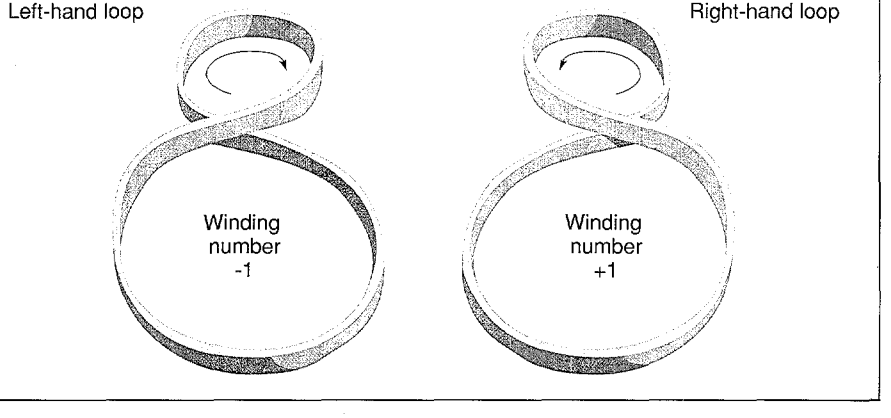
Figure 6. The Packaging of a DNA Loop

Artist's rendition shows the packaging of a DNA loop, first into the different orders of chromatin, and then into a twisted loop within a metaphase chromosome. The DNA's double helix makes two turns about a histone octamer to form the nucleosome, the repeating unit in chromatin. The chromatosomes, or nucleosomes bound to H1, are shown forming a thick chromatin fiber 34 nanometers in diameter. In this model, the supercoil of chromatosomes is further condensed into a twisted loop attached at a single point to the protein scaffold. The spiraling array of twisted loops constitutes the familiar metaphase chromosome, which is visible through the light microscope.

DNA loops is 50,000 base pairs, or 17 micrometers in length, each loop of DNA can form a string of nucleosomes that are either coiled to form 2.6 micrometers of a 10 nanometer fiber, or supercoiled into 0.4 micrometers of a 34 nanometer fiber. Thus, to create the thickness of a sister chromatid (Figure 1), which is 1 micron in diameter, would require just one more order of chromatin folding above the 34 nanometer supercoil. Figure 6 shows a possible model of this final level of chromatin folding.

How is the packaging of DNA loops controlled in response to chromosome functions? Evidence suggests that the inactive form of chromatin is the 34-nanometer supercoil or solenoid of nucleosomes. For both DNA transcription and genome replication this supercoil of nucleosomes must first be uncoiled to the linear array of nucleosomes and then the DNA must uncoil even further to allow access of the transcriptional machinery or the replication machinery to the DNA sequences. Whenever DNA is constrained by proteins to form a loop, DNA supercoiling becomes an important consideration in understanding DNA structure-function relationships. DNA supercoiling has been subjected to extensive experimental and mathematical analysis.

Figure 7. Configurations of a Closed Loop of Ribbon with Winding Numbers +1 and -1



Consider a model in which each DNA loop is firmly attached to the protein scaffold of a chromosome and is therefore somewhat analogous to a closed loop of ribbon. A closed loop of ribbon has a topologically invariant property known as the winding number, which is the number of twists in the ribbon plus the number of times the ribbon crosses itself, that is, coils about itself. The winding number is an integer or half-integer and remains constant unless the ribbon is cut. Each complete twist and each complete crossing adds $+1$ or -1 to the winding number depending on the direction of the twist or crossing. A right-handed twist (the same direction as the thread of a standard screw and the standard helical structure of a double-stranded DNA molecule) is positive, and a left-handed twist is negative. Similarly, a crossing that produces an extra right-handed loop in a loop of ribbon is positive, and a crossing that produces an extra left-handed loop in a loop of ribbon is negative (see Figure 7).

Now consider a loop of double-stranded DNA. Unconstrained DNA has 10.4 to 10.6 base pairs in each complete turn of the double helix. Taking the value 10.6 base pairs per helical turn, the twist (*Tw*) of a loop of unconstrained DNA consisting of N base pairs would be $N/10.6$. Because a double-stranded DNA molecule already has a helical structure, a loop of DNA further coiled about itself is said to be supercoiled. The linking number (*Lk*) of a closed loop of DNA is defined in terms of the twist and the number of supercoils, or writhe (*Wr*), through the equation $Lk = Tw + Wr$. Twists can be converted into supercoils, but *Lk* must remain constant in a DNA loop whose ends are fixed, in analogy with the constancy of the winding number of the loop of ribbon. If the loop is closed, the linking number must be an integer.

As an example, suppose three helical turns of a linear stretch of DNA are unwound and the ends are then joined. The linking-number change resulting from the unwinding is -3, and the loop can take on any of the three configurations shown in Figure 8. Moreover, the three configurations can be converted into one another without cutting the DNA. DNA configured as in (b) and (c) is said to be negatively supercoiled.

As shown in Figure 3, the DNA in the nucleosome core particle has 1.7 left-handed supercoils and in early studies it was expected that the linking-number change associated with the dissociation of a core particle would be -1.7. However, the experimentally determined linking-number change was -1.02. Although this difference was unexpected and initially controversial, it is easily explained by the change in twist between the DNA constrained in the core particle and free DNA in solution. The average DNA helical repeat on the core particle as measured from its crystal structure is 10.1 base pairs per turn. If we take the average helical repeat of free DNA as 10.6 base pairs per turn, the difference in twist between the DNA in the core particle and free DNA would be $146/10.1 - 146/10.6$ that is, 0.68. Thus the linking-number change associated with the core particle $\Delta Lk = -1.7 + 0.68 = -1.02$ as observed.

Now we can suggest how a DNA loop packaged as a 34-nanometer supercoil of nucleosomes (see Figure 6) could be unwound during interphase. If negative supercoils previously constrained by the nucleosomes are released, then negative supercoiling must be taken up by the linker DNA joining one nucleosome to another. This negative supercoiling would favor the unwinding of a 34-nanometer supercoil of nucleosomes. As suggested above, the acetylation of histones releases DNA that was negatively supercoiled about the histone octamer, presumably by unwinding DNA from the ends of the nucleosome.

The reverse process of chromosome condensation to the metaphase configuration (see Figure 1) requires that the 34-nanometer supercoil be further coiled into higher orders of coiling(s). Perhaps histone-H1 phosphorylation introduces additional supercoiling into a packaged DNA loop causing the higher order of coilings of metaphase chromosomes.

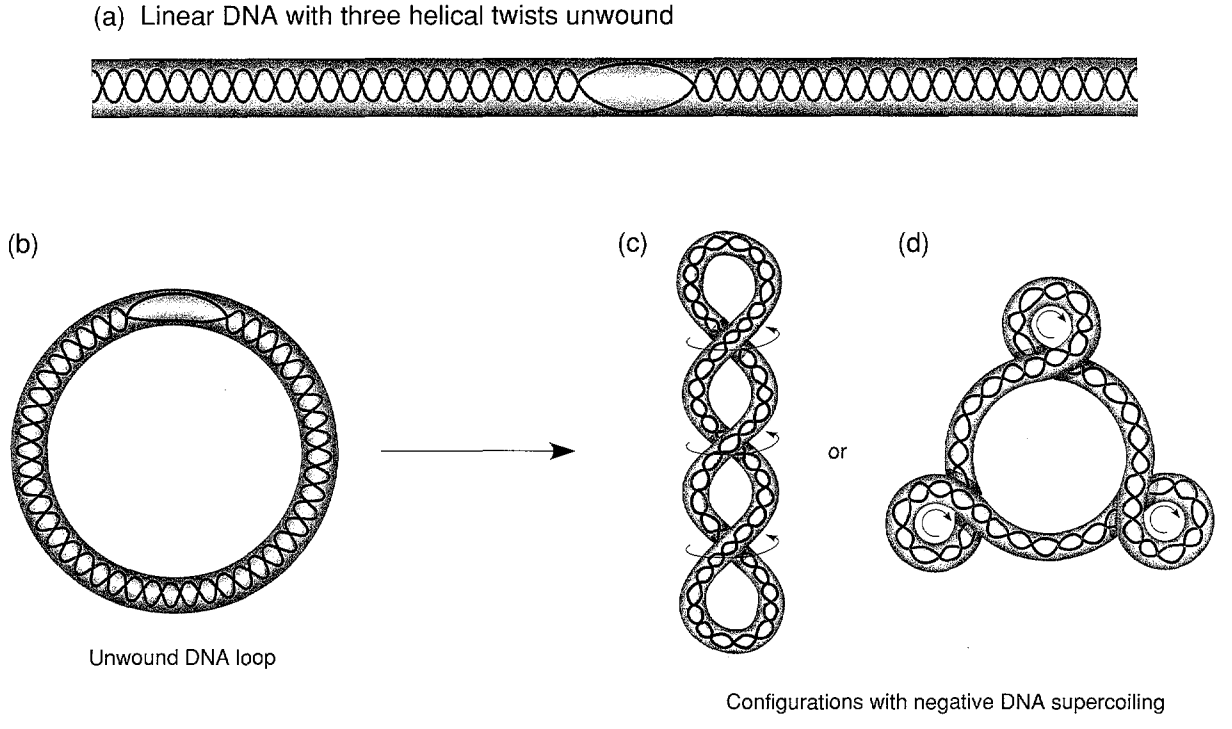


Figure 8. Negative Supercoiling of a Closed DNA Loop
If three helical twists of a linear, double-stranded DNA molecule are unwound as shown in (a) and the ends are then joined, the resulting DNA loop can take on the configurations shown in (b), (c), and (d). All three have the same linking number. In (b) the circular molecule is missing three helical twists that would be present in the normal structure. In (c) the three twists are restored and the loop forms a right-handed superhelix with three crossings in (d) the three twists are restored, but the loop forms three extra left-handed loops. Configurations (c) and (d) are referred to as negative DNA supercoiling.

Figure 8 shows in outline the different orders of packaging of DNA loops into the different orders of chromatin structure and into metaphase chromosomes. It appears that the reversible chemical modifications of acetylation and phosphorylation of histones are involved in the structural transitions undergone by a chromosome during the cell cycle. These structural transitions are dictated by the functional requirements of chromosomes.

Conclusion

Despite recent advances in understanding centromeres and telomeres, we are still a long way from understanding the relationships between structure and function of eukaryotic chromosomes. Relevant to this understanding will be the sequence information from the Human Genome Project. Although much interest is now focused on the mapping and sequencing of genes, the noncoding DNA regions clearly

contain information involved in the organization and functions of chromosomes. The constancy of the banding patterns of individual metaphase chromosomes reflects a highly reproducible pattern of long-range DNA folding, most probably directed by specific DNA-protein interactions and possibly by unusual DNA structures such as bent DNA segments. Superimposed on the very long-range order suggested by banding patterns is the packaging of the DNA loops by the histones together with other structural and regulatory proteins.

The existence of several subtypes of each histone raises the possibility that DNA loops containing different gene families could be packaged with different types of histones according to the requirements of the different cells. DNA control regions of active genes must be packaged in a fashion that makes them accessible to gene-regulating proteins, whereas regions containing permanently repressed genes of a particular cell type may be packaged so that they are inaccessible to such proteins. Such packaging may also determine the availability of DNA regions to chemical damage. Thus a knowledge of the organization of chromosomes is essential to an understanding of the central processes of cell differentiation and the orderly development of complex organisms as well as the processes of DNA damage in chromosomes. ■

Further Reading

E. Morton Bradbury, "Reversible Histone Modifications and the Chromosome Cell Cycle." *BioEssays*, Volume 14, No. 1. January 1992.

Morton Bradbury received a bachelor of science degree in physics and a Ph.D. in biophysics from King's College, University of London, in 1955 and 1958, respectively. After completing his postdoctoral research at Courtauld Research Laboratory, he was appointed head of the Department of Molecular Biology at Portsmouth (England) Polytechnic in 1962, where he remained until his appointment at UC Davis in 1979. He became leader of the Life Sciences Division at Los Alamos in 1988. Bradbury's research has been devoted to understanding whether chromosome organization and chromosome structure are involved in determining how a cell looks and behaves; the structure and function of active chromatin; and the process by which chromosomes condense prior to cell division. In pursuing his investigations, Bradbury has combined the results of measurements derived from the use of a wide range of techniques, including optical spectroscopy, nuclear magnetic resonance, x-ray diffraction, electron diffraction, and neutron diffraction. The recipient of numerous award and honors, Bradbury has also chaired a number of scientific organizations, including the British Biophysical Society, the International Council for Magnetic Resonance in Biology, and the Neutron/Biology Committee of the Institut Laue-Langevin. Bradbury is a member of HERAC and a member of the HERAC subcommittee on structural biology.