a Use the cord to reproduce the transformation shown in going from Fig. 6.1(b) to (c). For each full turn of 360° wound into one end of the cord, how many times does it cross over itself?

b Use the same cord to demonstrate that a left-handed toroidal supercoil, as in Fig. 6.2(a), transforms into a right-handed interwound coil, as in Fig. 6.2(b), when it is shaken out. To do this, first make a straight left-handed coil by wrapping the cord around a stick. (Hint: hold the stick and one end of the cord in your left hand, and wind the cord onto the stick using your right hand.) Then carefully remove the stick and bend the coil around to make a toroid; and finally hold the two ends together in one hand, and shake it out vertically. How many times does the right-handed interwound form cross over itself, for each turn of left-handed toroid around the stick?

6.2 For this exercise you should use the same cord as in Exercise 6.1, but now you must draw a black (or coloured) line (or stripe) along the entire length of the untwisted cord, in order to signify zero twist. To make a closed loop from the cord you can fasten the two ends together with sticky tape; or you can use a short dowel if you are working with rubber tubing; or you can hold the two ends together with your hand, as in Exercise 6.1.

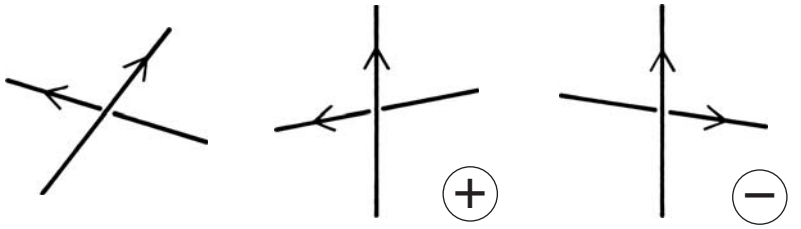
Reproduce the transformations shown in Figs 6.4 and 6.5, by winding into the cord three full turns of either right-handed or left-handed twist, to give Lk = +3 or −3, respectively. Which of the shapes (b) to (e) is the most stable for your model? Count the number of crossovers, and then use Figs 6.4(e) and 6.5(e) as guides to decide upon the sign of Wr, by looking at the handedness of the crossovers. Also, follow the path of the black stripe in order to identify the sign and magnitude of twist Tw. Confirm that Lk = Tw + Wr.

Untwist the cord fully and repeat with, say, Lk = +5 or −5.

6.3 This is an exercise on determining the sense (+ or −) with which an individual crossover contributes towards the writhe Wr.

First, identify some particular crossover in a picture of a writhed cord. Put an arrow on the upper segment of the cord (i.e. the one nearer to you), and follow the cord round until you reach the lower segment at the same crossover. Mark an arrow on this segment too, in the same sense as the upper arrow, when going along the contour-length of the cord.

The two crossed segments should now both be marked with arrows, as shown in (a), below. Next, rotate this local picture of the crossover until the arrow of the upper segment points towards the top of the page. If the lower arrow now points from right-to-left, as in (b), then the crossover counts as +1 to Wr; but if it points from left-to-right, as in (c), then the crossover counts as −1 to Wr.



Use this double-arrow test to confirm that there is negative Wr in each of the coils shown in Fig. 6.2; and also to confirm the signs of writhe Wr assigned to the coils shown in Figs 6.4(e) and 6.5(e).

(You can, equivalently, associate a 'right-hand rule' with the situation shown in (b); and if you are familiar with rules of this kind in electricity and magnetism you might well find it more useful. Thus, if you straighten the thumb of your right hand and bend the fingers round, and then point the thumb along the upper segment, your fingers will indicate the direction of the lower segment, for positive Wr. But if you try this with (c), you will find that the lower segment goes the other way: so there we have a negative contribution to Wr.)

6.4 For this exercise you can use either a cord with a painted stripe, as in Exercise 6.2, or else a ribbon or strip of paper which has been colored on one side with a felt-tipped pen, as in Fig. 6.6.

a Wrap a portion of the cord twice, in a left-handed sense, around an ordinary soft-drink or beer can. Use your two hands to represent the two blocks to which the ends of the cord are attached, as in Fig. 6.6(c). Then remove the can from the wrapped cord, and pull the ends apart as in Fig. 6.6(a). How many turns of twist Tw have been created by the wrapping? Are these new turns right-handed or left-handed?

b The vast majority of DNA in our chromosomes wraps about protein spools of the kind shown in Fig. 1.5. There, the DNA wraps twice in a left-handed sense around each spool, just as in our ribbon-around-the-beer-can exercise. But when the DNA is removed from such a protein spool, it turns out by experiment that only one turn of twist is created for each two turns of wrapping. In other words, Tw $= -1$ instead of -2 as would be expected from the simple models of part (a) and Fig. 6.6(a). By how many turns would the DNA have to be pre-twisted, before being wrapped on the spool (or else during the process of wrapping), to account for this result? Why might the protein spools be designed to lessen the expected change in twist Tw, on wrapping or unwrapping the DNA? (See Germond, J.E., Hirt, B., Oudet, P., Gross-Bellard, M., and Chambon, P. (1975)

Proceedings of the National Academy of Sciences, USA **72**, 1843–7 for the first experiment of this kind ever performed.)

6.5 Suppose that DNA wraps around a cylindrical protein into n complete turns of left-handed supercoil; for example, $n = 2$ in Fig. 6.6(c). If we represent the stress-free, relaxed DNA as a ribbon, we can easily find by direct experiment in such cases that $\mathrm{Lk} = -n$, where n is a positive, whole number of turns. Thus, Lk as so defined is necessarily an integer; and the negative sign of Lk corresponds to the left-handed sense of the wrapping.

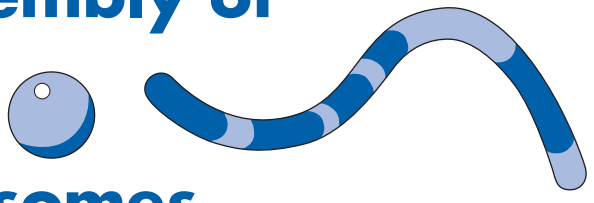
Now we can further describe the path of the coiled DNA in terms of a pitch angle α as shown previously in Fig. 5.4. The angle α is likewise negative for left-handed wrapping. Furthermore, by combining results from Chapters 5 and 6, it can be shown that:

$$\mathrm{Tw} = n \sin\alpha;\ \mathrm{Wr} = \mathrm{Lk} - \mathrm{Tw} = -n(1 + \sin\alpha)$$

For $n = 1$, calculate Tw and Wr for $\alpha = 0°$, $-30°$, $-60°$, and $-90°$.

CHAPTER 7

The Assembly of DNA into Chromosomes



In Chapter 1 we gave a general description of the biology of a typical cell, and we explained how DNA plays a central role in that biology, by specifying the construction of protein molecules. Since then, we have considered DNA mainly as a simple, double-helical thread which undergoes transcription and replication, and which wraps itself around protein spools. Our task in the present chapter will be to describe the assembly of DNA into chromosomes. How does DNA fold into the highly compact chromosomes shown in Fig. 1.3, when a cell is about to divide? And how does it organize itself at other stages in the life of a cell, as in Fig. 1.2, when it has to make the RNA that codes for protein? Our answers to these questions will be, unfortunately, less secure than we would prefer. Only a few aspects of chromosome structure are known with confidence, while the rest require a lot of educated guesswork.

We begin by describing in some detail the proteins which make up any histone spool, and the DNA that coils around them. Next we explain how a string of successive histone spools can coil into a '300 Å fiber', if the conditions are right. Then we describe how these 300 Å fibers might fold into a series of loops along some protein 'scaffold'; how genes might work along the hypothetical loops; and how the scaffold complete with loops might coil once again into the form of the compact chromosome which we can see with a microscope on cell division. Finally, we explain how chromosomes are assembled from proteins and DNA, at least at a rudimentary level; and how they might be disassembled when genes are activated for the synthesis of RNA.

Now you may recall from Fig. 1.1 that the DNA in our chromosomes is compacted by a factor of about 10 000 in total, as compared

with the length of a simple, double-helical thread. When DNA wraps around the histone spools, its overall length is reduced by a factor of about 6; so when it wraps into various other kinds of structure within a chromosome, its overall length must be reduced again by a further factor of 1500. Indeed, each individual chromosome contains a remarkably long length of DNA: typically 1 to 10 million base-pairs in yeast, or 50 to 400 million base-pairs in humans. If this DNA were not highly compacted in some definite way, then life as we know it could not exist. Throughout our presentation we shall consider, in a somewhat speculative way, how this degree of compaction might be accomplished.

Let us begin with a brief history of the subject. Chromosomes were studied for many years by geneticists, but not at a molecular level. Those scientists knew that chromosomes were made from a mixture of protein and DNA, but they did not know whether this mixture might possess any regular structure or organisation. In 1973, Dean Hewish and Leigh Burgoyne clarified the situation dramatically: they found that the majority of the DNA in a chromosome could be digested by a DNA-cutting enzyme into many small fragments of regular size, such as 200, 400, 600, 800, etc. base-pairs.

They were able to explain their result as follows. Suppose that chromosomes are made largely from a series of nearly identical units, each consisting of certain proteins in combination with 200 base-pairs of DNA. Now if the DNA-cutting enzyme were to act at every point where it found a weakness, perhaps in the regions of DNA which lie between units, then all of the long chromosomal DNA would be cut into short pieces of size 200 base-pairs. But if the cutting enzyme were to act at random, at only a limited number of points, then this long DNA would be reduced in size just to multiple units of 200, 400, 600, 800 base-pairs, etc., simply because not every point of weakness would be cut. Furthermore, if the individual units of 200 base-pairs were not of precisely determined size, but were to vary in size, for example from 180 to 220 base-pairs, then the lengths of the multiple units would vary also: from 360 to 440, 540 to 660, 720 to 880 base-pairs, etc. In such a case, these multiple units could hardly be seen as fragments of discrete size. But in fact, multiple units of discrete size are often seen up to DNA lengths as large as 2000 base-pairs in such experiments. Thus, one may conclude that certain proteins bind to this DNA in a rather precise fashion, once every 200 base-pairs, and thereby set its fundamental length.

It is now known that regular combinations of protein and DNA can be found once every 200 base-pairs along most of the length of DNA in any chromosome. These particles are called *nucleosomes*, and they are very important in biology. The fundamental length is

not always exactly 200 base-pairs in every kind of animal or plant, or in every type of tissue within a particular kind of animal or plant. In fact, these fundamental 'spacings' are known to vary from as short as 160 base-pairs in certain chromosomes, to as long as 260 base-pairs in others. So one suspects that there may be some variety in the kinds or numbers of protein that make up any nucleosome. We have already shown some crude pictures of nucleosomes elsewhere in the book, for example in Figs 1.5 and 6.6(c), without giving these particles a name, or explaining how they were made.

For the sake of accuracy, one should note that the repeating pattern seen by Hewish and Burgoyne in 1973 had first been seen 3 years earlier by Robert Williamson. He was studying mouse cells in tissue culture, and he saw DNA fragments of size 200, 400, 600, 800 base-pairs, etc., as products of degradation from cells that were not growing so well. But he thought that these fragments might come from the incomplete synthesis of long DNA molecules, as shown by previous work, rather than from the degradation of intact chromosomes. Because of this, his work attracted little attention; yet progress in chromosome research could have proceeded more rapidly, if more people had realized its implications.

Following the correct interpretation of these data by Hewish and Burgoyne, other workers in many laboratories across the world quickly provided more detailed information about the newly found nucleosome. For example, early in 1974 Ada and Donald Olins visualized a preparation of nucleosomes at high magnification using an electron microscope (a technique to be described in Chapter 9), and saw a series of protein 'beads' along an extended DNA 'string'. Such a result was clearly consistent with the idea of a regular, repeating structure for DNA in chromosomes. The next significant advance came later in 1974, with a report by Roger Kornberg and Jean Thomas on the identity and approximate organisation of the proteins in a typical nucleosome. They studied the physical properties of all the major chromosomal proteins, which are known as 'histones'; and they found that the majority of histone types, known as H2A, H2B, H3, and H4, could associate with one another in a stable fashion, so as to form a large protein particle around which the DNA could wrap. This particle is now known as the 'histone octamer', because it contains eight proteins in total: there are two copies of each of the four distinct kinds listed above. At first it was thought that all of the DNA in any 200-base-pair nucleosome might wrap solely about this histone octamer, but later it was realized that the histone octamer as such binds only to about 145 base-pairs of DNA. A ninth protein molecule, known as histone H1, was found to bind to the remaining 50 or so base-pairs. These results are summarized in a

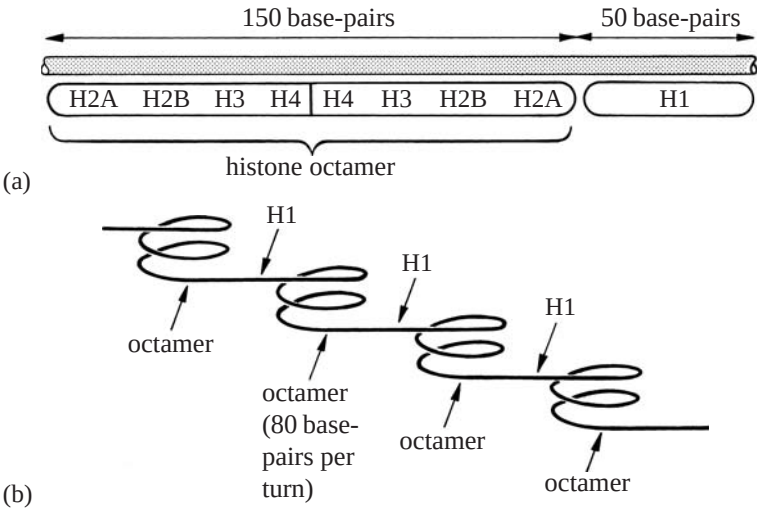


Figure 7.1 Histone proteins bound to DNA, and shown schematically in one dimension or three. The DNA is drawn as a long tube in (a), but as a long, curved string in (b).

simplified way in Fig. 7.1(a), which shows schematically the various histones strung out along the DNA in one dimension.

The three-dimensional structure of the histone octamer together with its DNA was not established firmly until 1977, when John Finch, Len Lutter and colleagues grew crystals of the histone octamer complete with DNA, that were suitable for analysis by X-ray diffraction methods. They also used electron microscopy to study the structure at low resolution, and enzyme-digestion methods to probe the structure in solution. By piecing together data gathered from all of these various methods (which will be described in Chapter 9), they were able to show that 145 base-pairs of DNA wrap almost twice around the histone octamer into a shallow, left-handed supercoil containing about 80 base-pairs per turn. Seven years later, in 1984, the same group with Tim Richmond succeeded in obtaining the structure of the histone octamer complete with its DNA to a resolution of 7 Å, solely by the method of X-ray diffraction. Finally in 1997, Richmond and colleagues analysed further the structure of the nucleosome core in better-ordered crystals at near-atomic resolution, thereby revealing many important details of how the DNA is wrapped on the surface of the histone octamer, as well as how the individual histone proteins fit together.

The old 1984 structure established that the diameter of the protein spool is about 60 Å, while the outer thickness of the DNA all around is about 20 Å, giving a diameter of about 100 Å for the particle overall. It also showed that the proteins H2A and H2B lie near both ends

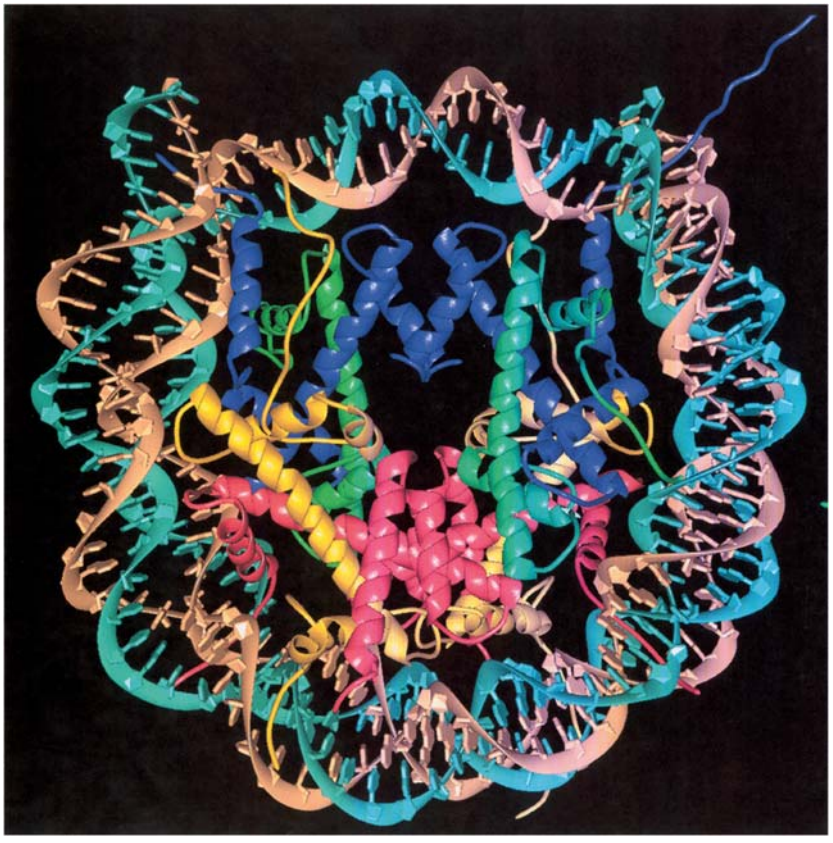


Figure 7.2 High-resolution X-ray diffraction structure of the complete nucleosome. The two-fold rotational symmetry of the protein about the vertical axis may be seen clearly. Colour-coding of proteins: H2A, orange; H2B, red; H3, blue and H4, green. The protein is mostly arranged as α-helices (see Chapter 8) and contacts between protein and DNA may be seen clearly. (The gaps between helices are filled by close-packed side-chains, not shown.) Courtesy of Tim Richmond.

of the DNA supercoil, while the proteins H3 and H4 lie near its center in a compact group. Hence the three-dimensional spool is a sort of wrapped-up version of the string of histones shown in Fig. 7.1(a). A model the new 1997 structure is shown in Fig. 7.2. Since the x, y, z coordinates in that new model are accurate to near-atomic resolution, many important aspects of nucleosome structure have been revealed that were not known before; for example: details of the protein fold and protein-to-protein contact; details of DNA roll, slide and twist; the locations of over 1000 water molecules; and the locations of cations such as magnesium or manganese.

The pictures shown in Figs 7.1 and 7.2 represent the simplest possible level of organisation for DNA and the proteins associated with

it in a chromosome. Let us now think about the ways by which a string of nucleosomes could fold into some higher-order structure. The regularity of the spacing – 200 base-pairs per nucleosome – should be an advantage for the assembly of some sort of ordered fiber, as opposed to an irregularly shaped 'clump'.

The structure of a polynucleosome fiber is still not known for certain, but one widely cited model is that of John Finch and Aaron Klug, later extended by those workers and Fritz Thoma and Theo Koller. In 1979 those scientists reported that one could stabilize the structure of a polynucleosome fiber by soaking it overnight in fixative, prior to taking pictures by electron microscopy. They then saw structures in the microscope that look like the model shown in Fig. 7.3. Unfortunately, the electron microscope does not give the degree of resolution that one would like to have, and so some interpretation is necessary. In their model, individual nucleosomes wrap into a

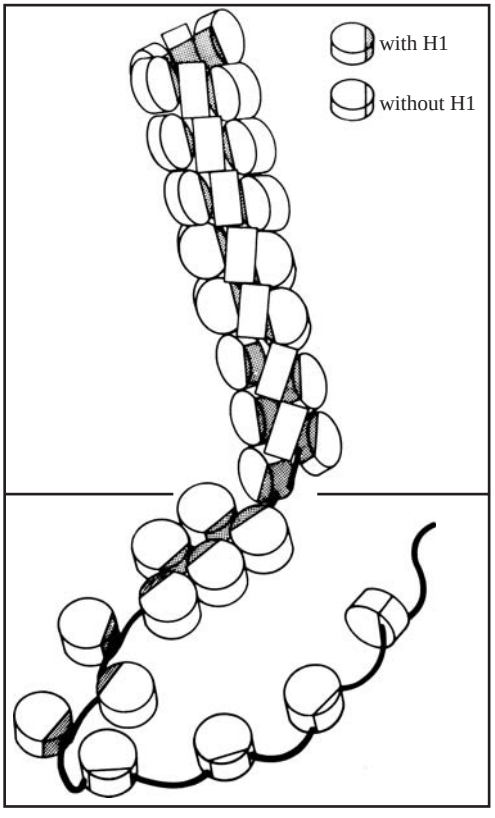


Figure 7.3 Histone octamers can be assembled onto long DNA like 'beads-on-a-string', and then these beads can wrap into some sort of more compact configuration, perhaps a flat spiral or '300 Å fiber' in the presence of histone H1. Adapted from F. Thoma *et al*. (1979).

compact spiral that advances by 110 Å per turn. Each turn of the structure contains about 5 to 6 nucleosomes, depending upon the solution conditions, and the DNA itself adopts a toroidal configuration. Histone H1 is thought to stabilize the folding of nucleosomes into this compact form: in the absence of histone H1, the nucleosomes are not 'frozen' by the fixative into some sort of spiral structure, but instead lie like 'beads-on-a-string' across the microscope support, as shown at the bottom of the picture. Whether or not histone H1 binds to the DNA depends on the chemical condition of the surrounding fluid, such as the amount of salt present.

The structure shown in Fig. 7.3 is known as a '300 Å fiber', because its outer diameter is typically 250 to 300 Å. The length of such a fiber is less than that of the constituent string of nucleosomes by a factor of about 6; so the length of the fiber is less than the length of a simple DNA thread by a factor of about 6×6, or approximately 40.

Histone H1 is located in the interior of this structure; but many other important aspects of the 300 Å fiber remain undetermined and are a subject of current research. Furthermore, due to the lack of precise structural definition in studies made by electron microscopy, various alternative models for the 300 Å fiber have been proposed: for example in the absence of fixative a highly irregular but coiled structure; or even an irregular zig-zag model, where the nucleosomes do not coil at all but follow a zig-zag path up the fiber axis.

Finally, we must emphasize that there are other abundant and important proteins in the cell nucleus as well as the histones, which serve to influence chromosome structure. In rank order after the histones, the next most abundant kinds of DNA-binding protein are the so-called 'high-mobility group' or 'HMG' proteins. These small proteins were noted by early investigators because they ran quickly through gels in electrophoresis experiments (see Chapter 9) when using an acetic acid–urea buffer. The three main types of HMG protein are known as HMGA, HMGB and HMGN. (They were known formerly as HMG-I,Y , HMG-1,2 and HMG-14,17 respectively.) Each class is about 1% to 10% as abundant as the histones in most tissues. All of these HMG proteins can associate with nucleosomes. HMGN is known to be associated with regions of DNA within actively transcribing genes, where the chromatin is partly unravelled; but the roles of HMGA and HMGB remain uncertain at present.

We have now summarized some of the more important small-scale features of chromosome structure. Let us therefore proceed to consider the large-scale features. Here the situation becomes very complicated. The problem is that, for most of the time, the DNA and histones are spread so uniformly in the cell nucleus that they show few distinguishing features, and so remain 'invisible' to both the