Further Reading

Allshire, R.C., Cranston, G., Gosden, J.R., Maule, J.C., Hastie, N.D., and Fantes, A.P. (1987) A fission yeast chromosome can replicate autonomously in mouse cells. *Cell* **50**, 391–403. The yeast–mouse hybrid chromosome having two different diameters (shown schematically in Fig. 7.7) came from cell-line F1.1 of this paper.

Bednar, J., Horowitz, R.A., Dubochet, J., and Woodcock, C.L. (1995) Chromatin conformation and salt-induced compaction: three-dimensional structural information from cryoelectron microscopy. *Journal of Cell Biology* **131**, 1365–76. Proposal of an irregular zig-zag model for the packing of nucleosomes in a 300 Å fiber.

Cai, S., Han, H.J., and Kohwi-Shigematsu, T. (2003) Tissue-specific nuclear architecture and gene expression regulated by SATB1. *Nature Genetics* **34**, 42–51. Possible organisation of nuclear sub-structures by the SATB1 protein.

Callan, H.G. (1982) Lampbrush chromosomes. *Proceedings of the Royal Society of London* **B 214**, 417–48. Many beautiful pictures of scaffolds and loops.

Cook, P.R. (1989) The nucleoskeleton and the topology of transcription. *European Journal of Biochemistry* **185**, 487–501. Are genes located on the outer parts of loops, or at their bases?

Davey, C., Pennings, S., Meersseman, G., Wess, T.J., and Allan, J. (1995) Periodicity of strong nucleosome positioning sites around the chicken adult beta-globin gene may encode regularly spaced chromatin. *Proceedings of the National Academy of Sciences, USA* **92**, 11210–4. Histone octamers bind to natural DNA sequences with a variation in affinity of 300-fold; such wide variations in affinity may help to locate nucleosomes in living cells.

Disney, J.E., Johnson, K.R., Magnuson, N.S., Sylvester, S.R., and Reeves, R. (1989) High mobility group protein HMG-I localizes to GQ and C bands of human and mouse chromosomes. *Journal of Cell Biology* **109**, 1975–82. The HMG-I protein (i.e. HMGA) is found mainly in inactive regions of a chromosome.

Farr, C.J., Bayne, R., Kipling, D., Mills, W., Critcher, R., and Cooke, H.J. (1995) Generation of a human X-derived mini-chromosome using telomere-associated chromosome fragmentation. *EMBO Journal* **14**, 5444–54. Early steps towards making an authentic mini-chromosome for human cells.

Hirano, T. and Mitchison, T.J. (1993) Topoisomerase II does not play a scaffolding role in the organisation of mitotic chromosomes assembled in Xenopus egg extracts. *Journal of Cell Biology* **120**, 601–12. Study of the large-scale structure of chromosomes regarding scaffolds and loops.

Iyer, V. and Struhl, K. (1995) Poly (dA.dT), a ubiquitous promoter element that stimulates transcription *via* its intrinsic DNA structure. *EMBO Journal* **14**, 2570–9. Long runs of either AA/TT or else GG/CC will stimulate transcription from a nearby gene in yeast, by removing nucleosomes near the gene so as to let transcription factors bind more readily.

Kim, J.H., Lane, W.S., and Reinberg, D. (2002) Human 'elongator' facilitates RNA polymerase II transcription through chromatin. *Proceedings of*

the National Academy of Sciences USA **99**, 1241–6. A large complex of proteins known as 'elongator' helps RNA polymerase to transcribe genes within chromatin, and also contains a histone acetylase enzyme.

Kurdistani, S.K. and Grunstein, M. (2003) Histone acetylation and deacetylation in yeast. *Nature Reviews in Molecular Cell Biology* **4**, 276–81. The roles of histone acetylation and de-acetylation in chromosome dynamics.

Lagarkova, M.A., Iarovaia, O.V., and Razin, S.V. (1995) Large-scale fragmentation of mammalian DNA in the course of apotopsis proceeds *via* excision of chromosomal DNA loops and their oligomers. *Journal of Biological Chemistry* **270**, 20239–41. Indirect evidence for loops of size 50 000 or 300 000 base-pairs in chromosomes, by gentle fragmentation on a large scale.

Li, Q. and Wrange, O. (1995) Accessibility of a glucocorticoid response element in a nucleosome depends on its rotational positioning. *Molecular and Cellular Biology* **15**, 4375–84. Transcription factors may bind with widely different affinities to the same DNA sequence in chromatin, depending on how such DNA curves about the histone proteins.

Rattner, J.B. and Lin, C.C. (1985) Radial loops and helical coils coexist in metaphase chromosomes. *Cell* **42**, 291–6. Pictures of duplicated metaphase chromosomes as made from a spiral rod.

Richmond, T.J. and Davey, C.A. (2003) The structure of DNA in the nucleosome core. *Nature* **423**, 145–50. Detailed analysis of DNA curvature in terms of roll-slide-twist or other helical parameters for almost two turns of left-handed DNA supercoil.

Sandaltzopoulos, R., Blank, T., and Becker, P.B. (1994) Transcriptional repression by nucleosomes but not H1 in reconstituted preblastoderm *Drosophila* chromatin. *EMBO Journal* **13**, 373–9. Spacing by histone H1 from 180 to 220 base-pairs in a fly-egg extract, and studies of gene function there.

Sivolob, A.V. and Khrapunov, S.N. (1995) Translational positioning of nucleosomes on DNA: the role of sequence-dependent isotropic DNA bending stiffness. *Journal of Molecular Biology* **247**, 918–31. A good model for rotational and translational positioning of histone octamers on different DNA sequences, based on how easily various sequences will bend.

Suka, N., Shinohara, Y., Saitoh, Y., Saitoh, H., Ohtomo, K., Harata, M., Shpigelman, E., and Mizuno, S. (1993) W-heterochromatin of chicken: its unusual DNA components, late replication, and chromatin structure. *Genetica* **88**, 93–105. Curved DNA makes up a large part of the chicken W sex chromosome.

Sumner, A.T. (1990) *Chromosome Banding* Unwin Hyman, London. An authoritative account of the methods used to study metaphase chromosomes.

Tremethick, D. (1994) High mobility group proteins 14 and 17 can space nucleosomal particles deficient in histones H2A and H2B, creating a template that is transcriptionally active. *Journal of Biological Chemistry* **269**, 28436–42. HMG-14,17 (i.e. HMGN) can space histone octamers from 145 to 165 base-pairs, while making the chromatin more active towards transcription; and also can space histone hexamers from 125 to 145 base-pairs. (see Fig. 7.12)

Truss, M., Bartsch, J., Schelbert, A., Hache, R., and Beato, M. (1995) Hormone induces binding of receptors and transcription factors to a rearranged nucleosome on the MMTV promoter in vivo. *EMBO Journal* **14**, 1737–51. A careful study of nucleosome locations at high resolution in living cells, before and after gene activation by transcription factors.

Tsukiyama, T., Becker, P.B., and Wu, C. (1994) ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. *Nature* **367**, 525–32. First direct demonstration of ATP-dependent nucleosome remodelling.

Urata, Y., Parmelee, S.J., Agard, D.A., and Sedat, J.W. (1995) A three-dimensional structural dissection of Drosophila polytene chromosomes. *Journal of Cell Biology* **131**, 279–95. A detailed study of the internal structures of fly polytene chromosomes, using light microscopy.

van Holde, K. (1988) *Chromatin* Springer-Verlag, New York. A major reference in the field, giving an overview of DNA and chromosomes in the 1980s.

Widom, J. (1996) Short-range order in two eukaryotic genomes: relation to chromosome structure. *Journal of Molecular Biology* **259**, 579–88. Periodicities of sequence near 10.2 base-pairs can be found in the total DNA of organisms such as bacteria, yeast or worm, especially for the dinucleotides AA/TT and GC/GC, as expected for the preferred bending of such DNA into tight curves.

Wolffe, A. (1995) *Chromatin* Academic Press, London. An important book citing much of the latest work on DNA and chromosomes.

Yoshida, M., Horinouchi, S., and Beppu, T. (1995) Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. *Bioessays* **17**, 423–30. Discovery of two chemicals that can inhibit the histone acetylase enzymes with high specificity.

Zinkowski, R.P., Meyne, J., and Brinkley, B.R. (1991) The centromere–kinetochore complex: a repeat subunit model. *Journal of Cell Biology* **113**, 1091–110. A centromere from a mammalian chromosome is shown to consist of many structural repeats, each of which attaches separately to microtubules on cell division.

Bibliography

Allfrey, V., Faulkner, R.M., and Mirsky, A.E. (1964) The acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proceedings of the National Academy of Sciences, USA* **51**, 786–94. First inference that the acetylation of histones might be correlated with transcription.

Arents, G., Burlingame, R.W., Wang, B.-C., Love, W.E., and Moudrianakis, E.N. (1991) The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proceedings of the National Academy of Sciences, USA* **88**, 10148–52. High resolution pictures of the histone proteins without the DNA.

Finch, J.T., Lutter, L.C., Rhodes, D., Brown, R.S., Rushton, B., Levitt, M., and Klug, A. (1977) Structure of nucleosome core particles of chromatin. *Nature* **269**, 29–36. First crystals of the nucleosome core particle, comprising the histone octamer plus 145 base-pairs of DNA, and their structural analysis at low resolution.

Hewish, D. and Burgoyne, L. (1973) Chromatin sub-structure: the digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochemical and Biophysical Research Communications* **52**, 504–10. Discovery of a regular spacing of histone proteins along the DNA in chromatin.

Jackson, D.A., Dickinson, P., and Cook, P.R. (1990) The size of chromatin loops in HeLa cells. *EMBO Journal* **9**, 567–71. The 'loop sizes' obtained on histone depletion depend strongly on the exact procedures used.

Kireeva, M.L., Walter, W., Tchernajenko, V., Bondarenko, V., Kashlev, M., and Studitsky, V.M. (2002) Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. *Molecular Cell* **9**, 451–2. How transcribing polymerase displaces an H2A/H2B dimer from the nucleosome in its path.

Kleinschmidt, J.A., Fortkamp, E., Krohne, G., Zentgraf, H., and Franke, W.W. (1985) Co-existence of two different types of soluble histone complexes in nuclei of *Xenopus laevis* oocytes. *Journal of Biological Chemistry* **260**, 1166–76. Independent discovery of N1 and nucleoplasmin by these workers.

Kornberg, R.D. and Thomas, J.O. (1974) Chromatin structure: oligomers of the histones; Kornberg, R.D. (1974) Chromatin structure: a repeating unit of histones and DNA. *Science* **184**, 865–8; 868–71. Proposal for a histone octamer as the fundamental unit of protein–DNA assembly in chromosomes.

Laskey, R.A., Mills, A.D., and Morris, N.R. (1977). Assembly of SV40 chromatin in a cell-free system from *Xenopus* eggs. *Cell* **10**, 237–43. Discovery of a cell extract that will make authentic pieces of a chromosome.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–60. The first near-atomic resolution crystal structure of the nucleosome core particle.

Olins, A.L. and Olins, D.E. (1974) Spheroid chromatin units (bodies). *Science* **183**, 330–1. First pictures of beads-on-a-string for histones and DNA in chromosomes.

Paranjape, S.M., Krumm, A., and Kadonaga, J.T. (1995) HMG-17 is a chromatin-specific transcriptional coactivator that increases the efficiency of transcription initiation. *Genes and Development* **9**, 1978–91. HMG-17 (i.e. HMGN) can increase the spacing of nucleosomes in a fly-egg extract from 160 to 175 base-pairs, while activating transcription, by increasing the rate at which new RNA chains are initiated.

Semeshin, V.F., Demakov, S.A., Perez-Alonso, M., and Zhimulev, I.F. (1990) Formation of interdiscs from DNA material of the P-element in *Drosophila* polytene chromosomes. *Genetika (USSR)* **26**, 448–56. The P-element DNA can make new bands or else split old ones.

Thoma, F., Koller, T., and Klug, A. (1979) Involvement of histone H1 in the organisation of the nucleosome and of the salt-dependent superstructures of chromatin. *Journal of Cell Biology* **83**, 403–27. Electron microscope study of long arrays of nucleosomes in a chemically fixed condition. The source of Fig. 7.3.

Tremethick, D.J. and Drew, H.R. (1993) High mobility group proteins 14 and 17 can space nucleosomes *in vitro*. *Journal of Biological Chemistry* **268**, 11389–93. Phosphorylated proteins HMG 14, 17 (i.e. HMGN) from human placenta can space nucleosomes from 145 to 165 base-pairs, just as for a spacing-fraction from the frog-egg extract that contains many HMG-like proteins, and depends on ATP (*via* a protein kinase) for its activity.

Williamson, R. (1970) Properties of rapidly labeled deoxyribonucleic acid fragments isolated from the cytoplasm of primary cultures of embryonic mouse liver cells. *Journal of Molecular Biology* **51**, 157–68. The 200-base-pair spaced pattern of chromatin, but not interpreted as such.

Exercises

7.1 A piece of double-helical DNA, 3010 base-pairs long, is bound in a test tube to a series of histone octamers in the manner of Figs 7.1, 7.2 and 7.3. Then an enzyme like the one used by Hewish and Burgoyne (1973) is added to a sample. Fine-mapping experiments show that the enzyme cuts only at certain locations along the length of the DNA molecule, which are known to be the empty spaces or 'linkers' between histone octamers. These locations, measured in terms of the distance from one end of the 3010-bp DNA, were found by experiment to be (in units of base-pairs): 220, 430, 670, 870, 1090, 1320, 1520, 1740, 1930, 2140, 2350, 2580, 2790.

a Make a table of all possible DNA fragment lengths, for fragments containing either 1, 2, 3, or 4 histone octamers. This table should have four columns. Column 1 should list all fragment lengths of size $220 - 0 = 220$, $430 - 220 = 210$, etc. base-pairs, and should contain 14 numbers. Column 2 should list fragment lengths of size $430 - 0 = 430$, $670 - 220 = 450$, etc., and should contain 13 numbers. Column 3 should list fragments such as $670 - 0 = 670$, $870 - 220 = 650$, and contain 12 numbers. Finally, column 4 should list sizes of $870 - 0 = 870$, $1090 - 220 = 870$, etc., and contain 11 numbers.

What is the mean fragment size in each of columns 1, 2, 3, 4, and its standard deviation?

b What is the mean spacing of histone octamers in this particular sample? Divide each mean fragment size by the column number 1, 2, 3, or 4 to get an optimal value.

7.2 The model of the histone octamer and its associated DNA, that is shown in Fig. 7.2, may be represented crudely as a protein cylinder of diameter 60 Å and height 50 Å, around which are wrapped two complete turns of DNA having a diameter of 20 Å.

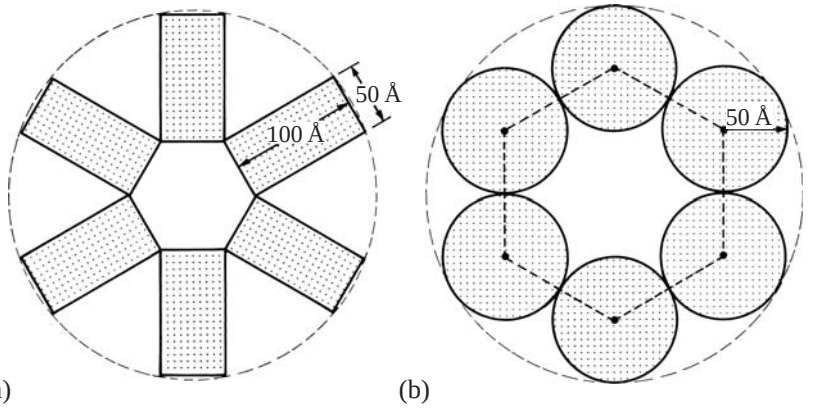
a Calculate the proportion of the total volume which is occupied by the DNA. For this rough calculation you may treat the DNA as two separate hoops, each with an inner diameter of 60 Å. Recall that the volume of a hoop is equal to the product of its cross-sectional area and its mid-circumference.

b Calculate the number of base-pairs in each hoop, assuming that the DNA has a length of 3.3 Å per base-pair.

7.3 The packing of nucleosomes into a usual model for the '300 Å fiber', as shown in Fig. 7.3, may be represented approximately by diagram (a), below. Here, to make the calculations especially simple, we have drawn the nucleosomes as closed rings of six, rather than as helical spirals with six nucleosomes per turn.

By considering each nucleosome as a cylinder of radius 50 Å and length 50 Å (for the protein plus DNA: see Exercise 7.2), estimate **a** the outer diameter of each ring, **b** the diameter of the inner hole, and **c** the height of the 300 Å fiber per ring of six nucleosomes.

A second possible packing scheme, also expressed in terms of closed rings, is shown in diagram (b) below. The nucleosomes are cylinders, just as before, but now their flat faces lie perpendicular to the axis of the 300 Å fiber, instead of lying parallel to this axis, as in diagram (a).



(a) (b)

Compute the dimensions **a**–**c** also for model (b); and compare them with the corresponding figures for model (a). (See Widom, J. and Klug, A. (1985) *Cell* **43**, 207–13, for evidence about key dimensions, and in particular that the axial spacing **c** is ≈110 Å.)

7.4 The telomeres of human chromosomes are made from a long, multiple repeat of the sequence (5′) TTAGGG (3′), as stated in the caption of Fig. 7.6. Often, such a repeat projects beyond the end of the double-helical DNA at each end of the chromosome, to leave a short, single strand of 6 unpaired bases as shown below:

(5′) …TTAGGGTTAGGGTTAGGG (3′)
(3′) …AATCCCAATCCC (5′)

Such an arrangement is known as a 'sticky end' of DNA. There are known to be many enzymes in the cell that can join or 'fuse' two such single-stranded ends of DNA molecules to one another, if they detect sufficient Watson–Crick base-pairing.

a Could two identical (5′) TTAGGG (3′) ends like the one shown above (with the second one found by rotating the first through 180° in the plane of the diagram) be joined by such an enzyme?

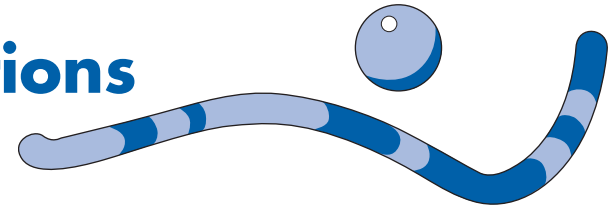
b Suppose that human telomeres were made not from repeats of TTAGGG, but from repeats of some other sequence such as TAGCTA or CGATCG. Would such hypothetical telomeres prevent the ends of chromosomes from joining to one another by Watson–Crick pairing? (See Ijdo, J.W., Baldini, A., Ward, D.C., Reeders, S.T., and Wells, R.A. (1991) *Proceedings of the National Academy of Sciences, USA* **88**, 9051–5, for a rare instance of telomere fusion.)

7.5 It is often possible to prepare, in the laboratory, a 'soup' or extract of proteins from a living cell which is able to synthesize some particular RNA molecule when a DNA 'template' molecule, having a specific sequence, is added to it.

If you study the process of transcription solely in this cell extract, on 'naked', histone-free DNA, what important features of the chromosome structure – features that could affect the activity of a DNA template in a living cell – would be excluded from your study? By scanning the diagrams of this chapter, identify at least four different structures of DNA within a chromosome that might influence gene activity in this way. (See Felsenfeld, G. (1992) *Nature* **355**, 219–24, for a survey of such concerns.)

CHAPTER 8

Specific DNA–Protein Interactions



In the previous chapter we explained how the vast majority of chromosomal DNA is associated with packaging proteins, such as histones. Those abundant proteins bind to most parts of the lengthy chromosomal DNA with only a slight preference for base sequence; and in organisms whose cells have nuclei, they serve mainly to compact the long DNA by a factor of about 10 000, into the minute volume of a cell nucleus. Chromosome compaction can control the access of RNA polymerase enzyme to the start site of genes (i.e., the 'promoter' sequences) and thereby affect the transcription of DNA into RNA. However, the study of packaging proteins alone cannot teach us how specific genes might be controlled, during the growth or development of cells into a mature organism, or during cellular response to environmental stimuli, or during the routine 'housekeeping' activities of fully developed cells.

For that reason, scientists have also studied the many less abundant DNA-binding proteins that bind specifically just to a small part of the long chromosomal DNA, in the proximity of the start sites for genes, or in other important locations. Some of these proteins can bind to DNA with high specificity for the base sequence, and can thereby carry out many specific biological tasks, such as control of transcription, repair of damaged DNA or unwinding of DNA. A few of these tasks are shown in Fig. 8.1.

The first DNA-binding proteins studied by scientists were regulatory proteins from bacteria, where they act to control perhaps the simplest genetic systems found in Nature. Many of these bacterial proteins act as '*repressors*' of gene activity (see the upper part of the picture) if they bind tightly to a base-sequence of DNA which overlaps the 'promoter' sequence, where an RNA polymerase enzyme

can also bind. They can thereby prevent the binding of RNA polymerase to a particular promoter, through direct competition for the same local segment of DNA. In general, such repressor proteins reduce the rate at which RNA is made from a promoter; and indeed such repression of RNA synthesis may be specific to just one or a few genes in an entire organism, if the repressor binds to only one or a few sites on an entire chromosome.

In bacteria, repressor proteins play an important role in reducing local rates of transcription; but in plants, animals and other organisms whose cells have nuclei – known collectively as eukaryotes[1] – the chromosome structure itself tends to repress transcription. Indeed, in nucleated organisms it is the activation of genes that seems to be the more important aspect of gene regulation. That process is managed by '*activators*' of transcription (see the middle part of Fig. 8.1) that bind specifically to DNA in the general vicinity of a binding site for RNA polymerase. The activator protein may then increase the rate at which RNA is made, by directly assisting the RNA polymerase enzyme and its auxiliary proteins to bind at the promoter sequence, through a network of protein-to-protein

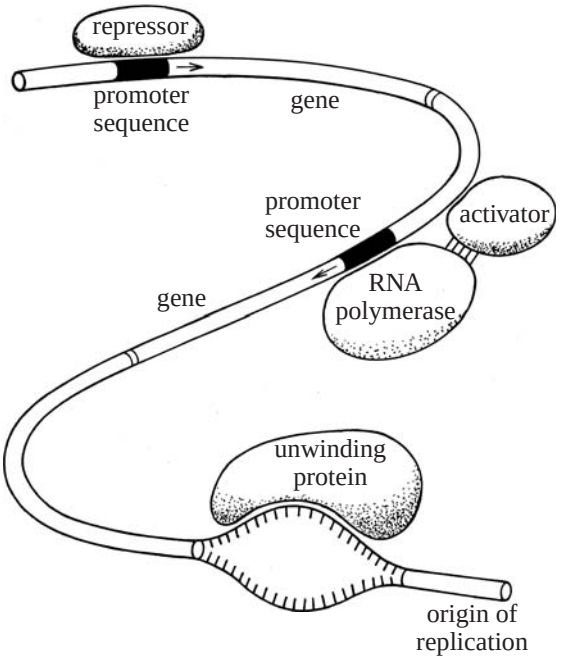


Figure 8.1 Proteins of various kinds may bind specifically to different DNA sequences, so as to carry out important biological tasks such as the repression or activation of individual genes, or unwinding the double helix in specific sites for copying or replication. A highly schematic picture, showing some of the DNA–protein interactions described in this chapter.

contacts; or else indirectly by helping to 'recruit' enzymes that can chemically modify the chromatin (see Chapter 7). For instance, certain transcription activators may direct histone acetylases to the general region of a specific gene. The resulting modification of histones may cause the chromatin to decompact near that promoter, and thereby make it more accessible to RNA polymerase and its auxiliary proteins.

'DNA looping' may represent a somewhat more complex example of how genes are regulated in three dimensions, and not just in one or two, by some linear or planar arrangement of DNA binding sites. In the latter case, two or more repressor or activator proteins may bind to the same piece of DNA, and then join together to create a small loop or coil, which can affect gene activity very strongly (either positively or negatively) on account of its stable structure.

We have mentioned so far repressors and activators and loops, but there are also many other kinds of regulatory protein that interact with the DNA, and so play important roles in the cell. Thus, a fourth kind of DNA-binding protein (see the lower part of Fig. 8.1) may help to unwind the DNA double helix near an 'origin of replication', where the duplication of old DNA into new DNA starts during every cell division. After such unwinding by these special kinds of protein, a DNA polymerase enzyme can read the bases on each separate strand by means of Watson–Crick pairing and then copy them; and so can make two new DNA double-helices from the two old strands. Also, proteins which bind to DNA may carry out many different biological tasks not shown here, such as the supercoiling of DNA, or the exchanging of DNA segments in 'recombination'. The cell also has elaborate assemblies of enzymes that recognize and repair DNA damage, which is a crucial aspect of cell viability.

In practice, all of the pictures shown in Fig. 8.1 are *greatly oversimplified*, since they omit so much important detail. For example, each specific interaction between protein and DNA must be highly complex from a chemical point of view, in order to distinguish one base sequence from others over the length of an entire chromosome. Also, each protein may be regulated in its ability to bind to specific sequences on the DNA, by a variety of factors including: (a) interaction with other proteins nearby; (b) any modification of the DNA such as base methylation (see Chapter 11); or (c) by binding to small molecules such as sugars and amino acids present in the cellular medium.

How can anyone hope to learn anything useful from such a cacophony of information, without making a big dictionary that would run to thousands of pages, and searching through the facts one at a time? The only hope for greater understanding seems to be