Yet scientists have recently developed other novel methods, which now permit the manipulation and detection of single DNA molecules in solution under near-physiological conditions. These single-molecule methods may be used, for example, to measure the local force needed to separate two strands of DNA; to study supercoiling of DNA when it is twisted; or to measure the forces required for packaging of DNA into the shell or 'capsid' of a bacterial virus.

Manipulation of single DNA molecules may be achieved with the help of micron-sized latex beads and micropipettes: the DNA is first attached to a latex bead, by a 'sticky tag' located at its very end, which then adheres to the surface of the bead. *Detection* may be achieved with the help of light microscopes and sensitive video-recorders, which can image the small beads (but not the DNA) and allow the path of the DNA to be deduced from the separation of the beads.

Several methods have been devised for applying small forces to the DNA. The earliest was to attach a double-stranded DNA molecule to the tip of a tiny glass needle, which then acts as an elastic cantilever beam as shown in Fig. 9.14 (a). Hence the force required to separate two strands of DNA may be deduced from the small elastic deflection of the glass needle, as it moves to the right in the picture. Another method is to attach double-stranded DNA to a tiny magnetic bead that is controlled by a magnet which can be rotated, as shown in Fig. 9.14(b). In that case, both tension as well as rotational torque may be applied to single molecules of DNA, allowing supercoiling to be studied.

Finally, the DNA may be attached to a tiny bead made from some material with high refractive index, such as latex. Then the bead can be held in place by intense laser light which has been focussed through a microscope lens: this is called an 'optical tweezer' or 'optical trap'. When a force is applied to the bead by tension in its attached DNA, the bead will be pulled elastically away from the laser light-focus, as if it were held there by an imaginary spring. Figure 9.14 (c) shows how this light-based method may be used to study the packaging of DNA into the shell or capsid of a bacterial virus.

Already scientists have measured these microscopic forces with sufficient accuracy to distinguish the separation of the two strands in A–T *versus* G–C rich DNA, and to study forces involved in binding DNA to protein. We expect that refinement of these techniques will have many other fruitful applications to DNA and DNA-protein interactions in the future.

Finally, we shall describe another technique that has been used in recent years to study the conformation of individual molecules

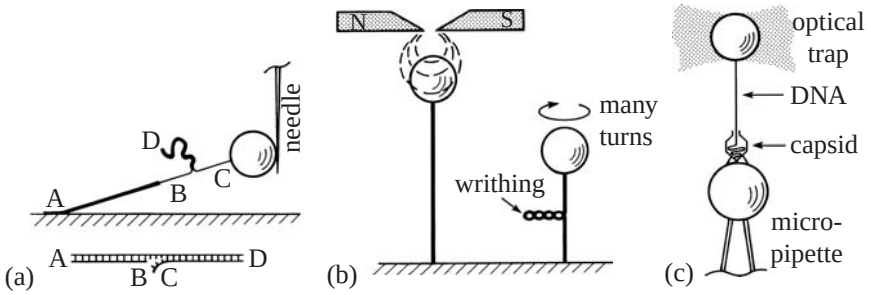


Figure 9.14 (a) Unzipping a single molecule of DNA. Double-stranded DNA with a break in one strand (lower diagram) was attached at end A to a microscope slide, but at end C to a small latex bead, which in turn was attached to a tiny glass needle. When that needle is moved to the right, the DNA becomes separated into two strands between B and C. The force is measured *via* the tip-deflection of the needle. (b) DNA attached to a magnetic bead can be twisted by means of a rotating magnetic field and made to writhe, if the tension imposed is sufficiently small – typically a fraction of 1 pico-newton (10^{-12} newton). On the left is the relaxed DNA, while on the right, after many turns, a portion of the DNA has writhed, as in Fig. 6.1(c). The length of DNA is about 10 000 base-pairs. (c) Packaging of DNA into a viral capsid. The capsid is attached to a bead held by a microscopic pipette, while one end of the DNA is attached to a bead held in an optical trap. The setup can be arranged to provide a constant force; and the rate at which the DNA is pulled by the tiny motor at the neck of the capsid can be measured. The length of the viral DNA is about 20 000 base-pairs.

attached to mica: Atomic Force Microscopy (AFM). In principle, this method involves dragging a hard, sharp probe over a flat specimen in a 'raster' of closely-spaced parallel straight lines in the x, y plane, and measuring the height z to which the probe must be raised in order to clear the object. Then a computer can process the measured $z(x, y)$ into a colour-coded contour map of the surface. In practice, the sharp probe – usually of silicon nitride – is mounted at the end of a fixed, elastic cantilever (of length about 50 μm); and the specimen is moved under it in three dimensions by piezo-drivers controlled by a computer. The specimen may be up to about 0.5 mm square, but normally much smaller areas are scanned.

AFM was first used to assay the variation of adhesive forces over the surface planes of metal or mineral crystals – hence the name; and it could also produce topographical contour maps of the surfaces, showing patterns of individual atoms. Assays of surface forces were achieved by measuring – with one of several possible methods – the minute deflection of the tip of the calibrated cantilever. But even if those forces are not of interest in a particular application, the deflection of the cantilever is still needed, as part of the control-loop for the z-direction movement of the specimen.

When AFM is used to visualize, say, the path of an individual DNA molecule attached to a mica surface, a number of special problems, which are not present with 'hard' specimens, have to be addressed. Thus, there is a tendency for the passing probe to dislodge the specimen, unless it has been stuck down by judicious application of magnesium to the mica base – where it forms a di-valent 'bridge' between the negatively charged DNA and the negatively charged mica. And in order to avoid damage to the 'soft' specimen, the sharp probe is usually oscillated vertically as it traverses, in so-called 'tapping mode', so that it taps across the specimen, something like a blind person tapping the ground with a stick. This procedure reduces lateral forces between the probe and the specimen, since the probe is in contact with the specimen for only a small fraction of the time.

The resolution that can be achieved with AFM on soft specimens such as protein and DNA typically can be as good as 1 Å vertically and 5 Å horizontally. A study of the wrapping of DNA around RNA polymerase using AFM and biochemical methods was mentioned in Chapter 6.

Other typical applications include the study of persistence length (a measure of elastic flexural stiffness) by using statistical analysis on the observed profiles of hundreds of DNA specimens; and investigation of the mode of action of restriction enzymes that break the DNA into pieces. An advantage of AFM is that it is equally effective on wet or dry specimens.

Note

1. See Appendix 1.

Further Reading

Brower-Toland, B.D., Smith, C.L., Yeh, R.C., Lis, J.T., Peterson, C.L., and Wang, M.D. (2002) Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA. *Proceedings of the National Academy of Sciences, USA* **99**, 1960–5. The use of 'optical tweezers' to study the unravelling of single-molecule DNA from nucleosomes.

Berge, T., Jenkins, N.S., Hopkirk, R.B., Waring, M.J., Edwardson, J.M., and Henderson, R.M. (2002) Structural perturbations in DNA caused by bis-intercalation of ditercalinium visualised by atomic force microscopy. *Nucleic Acids Research* **30**, 2980–6. Use of atomic force microscopy to study supercoiled DNA-drug complexes.

Bustamante, C., Bryant, Z., and Smith, S.B. (2003) Ten years of tension: single-molecule DNA mechanics. *Nature* **421**, 423–7. An overview of the application of single-molecule methods to study the mechanical properties of DNA.

Calladine, C.R., Collis, C.M., Drew, H.R., and Mott, M.R. (1991) A study of electrophoretic mobility of DNA in agarose and polyacrylamide gels. *Journal of Molecular Biology* **221**, 981–1005. Application of the ideas of A.G. Ogston to DNA gel-running. Source of the photographs in Fig. 5.5.

Diekmann, S. and Wang, J.C. (1985) On the sequence determinants and flexibility of the kinetoplast DNA fragment with abnormal gel electrophoretic mobilities. *Journal of Molecular Biology* **186**, 1–11. The kinetoplast DNA curves naturally in a plane, giving Wr = 0; but it gives non-zero Wr when distorted by supercoiling, due to its curvature of about 1.3° per base-pair.

Drew, H.R. (1984) Structural specificities of five commonly-used DNA nucleases. *Journal of Molecular Biology* **176**, 535–7. Source of the S1 nuclease gel shown in Fig. 9.9.

Hagerman, P.J. (1985) Sequence-dependence of the curvature of DNA: a test of the phasing hypothesis. *Biochemistry* **24**, 7033–7. First proof that the slow gel-running of certain DNA sequences is due to their curvature; and a deduction that these sequences proceed through the gel as superhelices of definite shape.

James, T.L. (ed.) (1995) Nuclear magnetic resonance and nucleic acids. *Methods in Enzymology* **261**, 1–640. A comprehensive series of articles, describing current progress on the use of NMR to study DNA and RNA.

Keller, W. (1975) Determination of the number of superhelical turns in simian virus 40 DNA by gel electrophoresis. *Proceedings of the National Academy of Sciences, USA* **72**, 4876–80. Gel-running of supercoiled DNA in the absence or presence of ethidium bromide, an intercalating dye.

Kozulic, B. (1995) Models of gel electrophoresis. *Analytical Biochemistry* **231**, 1–12. An excellent, knowledgeable and critical review of various models for DNA gel motion.

Leuba, S.H., Karymov, M.A., Tomschik, M., Ramjit, R., Smith, P., and Zlatanova, J. (2003) Assembly of single chromatin fibers depends on the tension in the DNA molecule: magnetic tweezers study. *Proceedings of the National Academy of Sciences, USA* **100**, 495–500.

Low, C.M.L., Drew, H.R., and Waring, M.J. (1984) Sequence-specific binding of echinomycin to DNA: evidence for conformational changes affecting flanking sequences. *Nucleic Acids Research* **12**, 4865–79. Source of the footprinting gel shown in Fig. 9.8.

Marini, J.C., Levene, S.D., Crothers, D.M., and Englund, P.T. (1982) Bent helical structure in kinetoplast DNA. *Proceedings of the National Academy of Sciences, USA* **79**, 7664–8. First indication that the slow gel-running of certain DNA sequences might be because they are curved.

Ogston, A.G. (1958) The spaces in a uniform random suspension of fibers. *Transactions of the Faraday Society* **54**, 1754–7. The classic work on the application of random statistics to gel-running phenomena, on which almost all later studies are based.

Rivetti, C., Guthold, M., and Bustamante, C. (1999) Wrapping of DNA around the *E. coli* RNA polymerase open-promoter complex. *EMBO Journal* **18**, 4464–75. Use of atomic force microscopy to map protein/DNA conformations.

Rodbard, D. and Chrambach, A. (1971) Estimation of molecular radius, free mobility and valence using polyacrylamide gel electrophoresis. *Analytical Biochemistry* **40**, 95–134. Application of Ogston's work to protein gel-running.

Sayre, A. (1975) *Rosalind Franklin and DNA*. W.W. Norton and Company, New York. The story of the discovery of a double-helical structure for DNA, told from the perspective of Rosalind Franklin, who collected the critical X-ray data from fibers.

Smith, S.B., Heller, C., and Bustamante, C. (1991) Model and computer simulations of the motion of DNA molecules during pulsed-field gel electrophoresis. *Biochemistry* **30**, 5264–74. A good theoretical study of alternating-field electrophoresis.

Whytock, S. and Finch, J.T. (1991) Substructure of agarose gels as prepared for electrophoresis. *Biopolymers* **31**, 1025–8. Source of Fig. 9.10.

Zimm, B.H. and Lumpkin, O. (1993) Reptation of a polymer chain in an irregular matrix: diffusion and electrophoresis. *Macromolecules* **26**, 226–34. A reptation theory for DNA motion through gels, that includes friction between DNA and the gel fibers.

Bibliography

Avery, O.T., MacLeod, C.M., and McCarty, M. (1944) Induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* Type III. *Journal of Experimental Medicine* **79**, 137–58. The classic work implicating DNA as the substance of genes.

Bernal, J.D. (1967) *The Origin of Life*. Weidenfeld & Nicolson, London. A good discussion of pre-biotic evolution and structures in biology, by one of the founders of the science.

Crawford, J.L., Kolpak, F.J., Wang, A.H.-J., Quigley, G.J., van Boom, J.H., van der Marel, G., and Rich, A. (1980) The tetramer d(CpGpCpG) crystallises as a left-handed double helix. *Proceedings of the National Academy of Sciences, USA* **77**, 4016–20. The sequence CGCG in a low-salt crystal form.

Dekker, N.H., Rybenkov, V.V., Duguet, M., Crisona, N.J., Cozzarelli, N., Bensimon, D., and Croquette, V. (2002) The mechanism of type IA topoisomerases. *Proceedings of the National Academy of Sciences, USA* **99**, 12126–31. Use of an attached magnetic bead to produce writhing in supercoiled DNA under tension. Source of Fig. 9.14(b).

Drew, H.R., Takano, T., Tanaka, S., Itakura, K., and Dickerson, R.E. (1980) High-salt d(CpGpCpG), a left-handed Z' DNA double helix. *Nature* **286**, 567–73. The sequence CGCG also goes left-handed in a high-salt crystal form.

Essevaz-Roulet, B., Bockelmann, U., and Heslot, F. (1997) Mechanical separation of the complementary strands of DNA. *Proceedings of the National Academy of Sciences, USA* **94**, 11935–40. Use of a flexible glass needle to measure the force required to separate the strands of double-helical DNA. Source of Fig. 9.14(a).

Franklin, R.E. and Gosling, R.G. (1953) Structure of sodium thymonucleate fibres: importance of water content. *Acta Crystallographica* **6**, 673–7. First identification of the 'B' and 'A' types of DNA in fibers.

Fuller, W., Wilkins, M.H.F., Wilson, H.R., Hamilton, L.D., and Arnott, S. (1965) The molecular configuration of deoxyribonucleic acid: X-ray diffraction study of the 'A' form. *Journal of Molecular Biology* **12**, 60–80. A refined model for the 'A' form of DNA.

Itakura, K. and Riggs, A.D. (1980) Chemical DNA synthesis and recombinant DNA studies. *Science* **209**, 1401–5. An early, important survey of the advances made in biology by the chemical synthesis of DNA.

Kozulic, B. (1994) On the 'door-corridor' model of gel electrophoresis. *Applied and Theoretical Electrophoresis* **4**, 117–59. An improved two-phase model for the motion of macromolecules through gels, which includes deformation of the gel by a migrating molecule.

Langridge, R., Marvin, D.A., Seeds, W.E., Wilson, H.R., Hooper, C.W., Wilkins, M.H.F., and Hamilton, L.D. (1960) The molecular configuration of deoxyribonucleic acid: molecular models and their Fourier transforms. *Journal of Molecular Biology* **2**, 38–64. A refined model for the 'B' form of DNA.

Leslie, A.G.W., Arnott, S., Chandrasekaran, R., and Ratliff, R.L. (1980) Polymorphism of DNA double helices. *Journal of Molecular Biology* **143**, 49–72. The wide variety of double-helical forms 'A, B, C, D, E' and their variants as seen by fiber X-ray diffraction.

Patel, D.J., Pardi, A., and Itakura, K. (1982) DNA conformation, dynamics and interactions in solution. *Science* **216**, 581–90. A review of some of the earliest work by NMR on large synthetic DNA. Source of parts of Fig. 9.4.

Pohl, F.M. and Jovin, T.M. (1972) Salt-induced cooperative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly (dG-dC). *Journal of Molecular Biology* **67**, 375–96. The left-to-right-handed transition of DNA in solution.

Schwarz, D.C. and Cantor, C.R. (1984) Separation of yeast chromosomal-sized DNAs by pulsed-field gradient gel electrophoresis. *Cell* **37**, 67–75. The discovery that alternating electric fields can separate very long DNA in gels according to size.

Smith, D.E., Tans, S.J., Smith, S.B., Grimes, S., Anderson, D.L., and Bustamante, C. (2001) The bacteriophage ϕ29 portal motor can package DNA against a large internal force. *Nature* **413**, 748–52. A study demonstrating how optical tweezers may be used to measure the force required to compact DNA into a bacteriophage capsid. Source of Fig. 9.14(c).

Viswamitra, M.A., Kennard, O., Jones, P.G., Sheldrick, G.M., Salisbury, S., Falvello, L., and Shakked, Z. (1978) DNA double-helical fragment at

atomic resolution. *Nature* **273**, 687–8. The crystal structure of ATAT, which 'melts' at the central TA step.

Wang, A.H.-J., Quigley, G.J., Kolpak, F., Crawford, J.L., van Boom, J.H., van der Marel, G., and Rich, A. (1979) Molecular structure of a left-handed double-helical DNA fragment at atomic resolution. *Nature* **282**, 680–6. First visualisation of left-handed DNA in a sequence CGCGCG.

Wüthrich, K. (1995) *NMR in Structural Biology*. World Scientific Publishing, Singapore. Methods for analyzing structures of biological molecules in solution.

Exercises

9.1 In an electron-density map such as that shown in Fig. 9.3, which of the atoms should scatter X-rays most strongly?

- **a** Rank in order of scattering power, from highest to lowest: carbon, phosphorus, oxygen, hydrogen, nitrogen.
- **b** Which part of the electron-density map in Fig. 9.3 was produced by the strongest scattering of X-rays? Use Figs 2.8(b) and 2.11(a) to identify the atom types.
- **c** Why are the hydrogen atoms in the Watson–Crick base-pairs not seen in Fig. 9.3?

9.2 Referring to Fig. 9.4, how many NMR peaks would you expect to find in the spectral region shown, and at –5°C, for double-helical DNA molecules which are specified by the following single-strand sequences?

- **a** (5′) CGCAATTGCG (3′)
- **b** (5′) AGCATGCATGCT (3′)
- **c** (5′) AGCATGCGCG (3′)

In each case, first construct the two-stranded version of the molecule, and then look for two-fold symmetry that may make some peaks equivalent to others.

9.3 Suppose we have a DNA molecule of length 1000 base-pairs, which contains three cutting sites, GAATTC, for a particular restriction enzyme; and suppose that these sites are located at 100, 350, and 550 base-pairs, respectively, from one end.

- **a** On complete digestion of the DNA by the restriction enzyme, how many kinds of DNA fragment will be produced, and of what size?
- **b** Which fragment will run fastest in an ordinary electrophoretic gel, and which will run slowest?

9.4a Suppose that you have the DNA from a virus, consisting of 100 000 base-pairs, and you wish to determine its complete nucleotide sequence. Find, approximately, the smallest

number of fragments that you could sequence individually, by use of present-day gel technology, in order to complete this task. Don't count the many overlapping fragments which would be needed to align partial sequences.

b A typical human cell contains 6×10^9 base-pairs of DNA, located on 23 pairs of chromosomes. What is the smallest number of fragments which you could sequence individually, in order to carry out the gigantic task of sequencing the complete DNA from a human cell? Again, don't count overlaps.

9.5 Figure 9.12(a) shows a flat disc migrating across a plane containing randomly-spaced point obstacles: bigger discs will contact more obstacles, and hence go through the gel more slowly. Similar concerns govern the motion of DNA through a gel in three dimensions. As an overall rule, one expects that flat discs in two dimensions with the largest areas, or DNA cylinders in three dimensions with the largest volumes, will contact the most gel fibers and hence go most slowly.

One can estimate the volume of the DNA cylinder for a straight piece of DNA, by modeling it as a cylinder of radius 10 Å and length 3.3 Å per base-pair. But for a curved, superhelical piece of DNA, one has to consider the volume of a 'circumscribing' cylinder into which the superhelix can just fit.

Find the apparent volume for one superhelical turn of curved DNA with the repeating sequence A_6N_4, using the following parameters for its shape: superhelical radius $r = 18.1 \times 3.3$ Å, contour length $N^* = 195.4 \times 3.3$ Å, and pitch angle $\alpha = -54.5°$ (see Table 5.1). First, find the pitch height p from $N^*\sin\alpha$ as shown in Figure 5.4, and confirm that $2\pi r$ equals $N^*\cos\alpha$. Then add 10 Å to the superhelical radius r, in order to account for half the thickness of the DNA itself. Finally, get the volume of the circumscribing cylinder from $\pi(r + 10)^2 p$.

Give your result as the ratio of two volumes, for curved DNA *versus* straight DNA of the same length N^*. Now repeat the same calculation, but using other curved sequences from Table 5.1, i.e. A_6N_2, A_6N_3, A_6N_5, A_6N_6, and A_6N_7. Which of these DNA sequences will go the most slowly through a gel, and which the most rapidly?

9.6 As explained in the text, closed circular DNA molecules run at a variety of speeds in a gel. That is because different linking numbers Lk can produce different interwound shapes, as shown, for example, in Figs 6.4(e) and 6.5(e) for Lk = ±3 or in Figs 6.4(a) and 6.5(a) for Lk = 0. Now it turns out that specimens with an open circular form (Wr = 0) run most slowly through a gel, while specimens

with higher (whether positive or negative) values of Wr run more rapidly, since they are more compact and so present a smaller effective volume to the interfering gel fibers which retard forward motion.

This pattern of behavior enables one to pick out specimens having Lk = 0 from a gel; but one cannot easily pick out a specimen having, say, Lk = +3 from one with Lk = −3, because they both run at very similar speeds. Also, one cannot easily distinguish large values of Lk, say +19 and +20, from each other, because those molecules are not sufficiently different in shape to run at significantly different speeds.

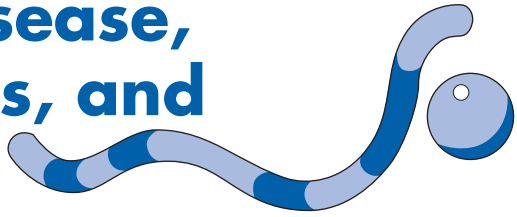
Scientists get around these problems by a technique which involves the addition of ethidium bromide or some other dye (such as chloroquine phosphate) to the gel, so that in any given specimen the DNA untwists, thereby acquiring a more positive value of Wr, since Lk does not change.

Given that the addition of a certain amount of ethidium bromide imparts exactly Tw = −12 to all molecules in a sample, predict the values of Lk and Wr for the slowest-running specimen in the range Lk = 0 to −20. Which will be the fastest-running specimen?

(Use the equation on p. 122, and assume that the slowest molecule always has Wr = 0, while the fastest molecule has maximal Wr, whether positive or negative.)

CHAPTER 10

DNA in Disease, Diagnostics, and Medicine



In the previous chapters, we have explained as simply as possible how DNA works in biology, according to various aspects of its three-dimensional structure. For example, the local pairing of bases in a Watson–Crick fashion across the double helix, as A with T or G with C, allows a DNA molecule to be copied from generation to generation, providing the general mechanism of inheritance for all life on Earth. Also, we have seen repeatedly how the structural properties of DNA as conferred by the sequence itself may be used in biology. For instance, the preferred unwinding of a double helix at sequences of the kind TATA defines many (but not all) of the start-sites for making more DNA or RNA, during replication or transcription. Again, the preferred bending of a double helix into the minor groove at certain sequences (e.g. AAA/TTT), or into the major groove at others (e.g. GGC/GCC), encourages the precise packaging of DNA as it curves around histone proteins in chromosomes, or as it curves about repressor and activator proteins that control transcription.

There is still much more to learn about DNA, than what we have explained so far. Unfortunately, however, not everything about the workings of DNA in biology is currently understood, especially in complex organisms such as humans, animals or plants. So while we cannot provide a comprehensive, authoritative account of how DNA works in biology, we can offer a snapshot of current understanding in a field that is growing rapidly. If you could sleep for a hundred years, like Rip van Winkle, you might awake to a world in which DNA and biology were understood completely and deeply, and you might find that this knowledge was being applied in ways that are unimaginable at the moment.