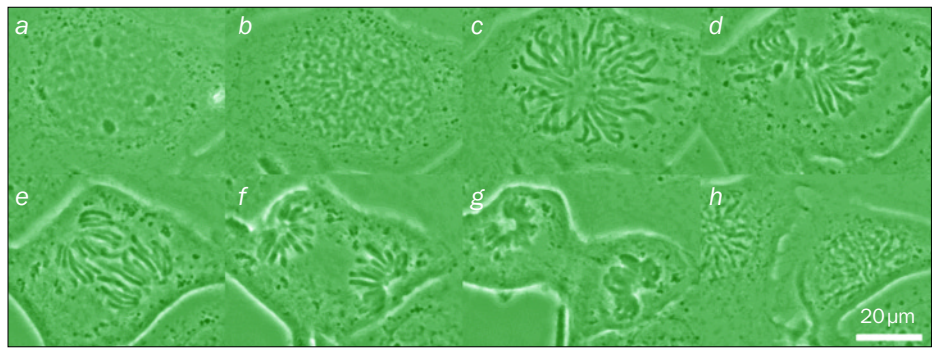
Advances in manipulating single DNA molecules are opening up new ways to study the micromachines that read, replicate and repair the double helix

The micromechanics of DNA

John F Marko and Simona Cocco

ALL LIVING cells contain long DNA molecules that carry genetic information in their base sequences. These molecules must be read so that proteins can be made, and they must also be duplicated and separated from one another so that cells can divide. Furthermore, if a cell's DNA suffers damage, it must be repaired. These essential processes are carried out by enzymes – proteins that catalyse biochemical reactions – that "walk" along the DNA double helix. So where does physics come in?

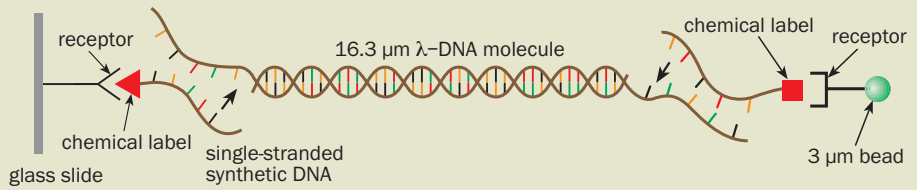
This cellular machinery must be considered from a biomechanical as well as a biochemical point of view, and this requires us to understand the micromechanical properties of DNA itself. Furthermore, since these DNA-processing machines operate at the nanometre scale with energies that are comparable to single thermal fluctuations, statistical mechanics is a key tool in analysing their behaviour. The combination of biochemistry with physical concepts and instrumentation is leading to new insights into how DNA is organized and processed inside cells.



MICHAEL POIRIER, LOUIS PASTEUR UNIVERSITY, STRASBOURG

The mechanics of life – the remarkable process of cell division is the result of the mechanical work done by nano-scale proteins on gigantic DNA molecules. Parts (*a*) to (*h*) illustrate how the organization of a cell's DNA changes drastically as the cell grows and divides. Paired, replicated chromosomes that contain about 1 m of DNA gradually become visible as short strings (*b*) and (*c*), before the copies are segregated to the daughter cells (*e*) and (*f*). A large number of enzymes play a critical role in this process, exerting forces of about 1 nN on the chromosomes. These images are phase-contrast light micrographs of cells from a species of newt with particularly large chromosomes.

1 Single-molecule DNA experiment



A large single molecule of double-stranded DNA can be manipulated by attaching short, single-stranded DNA molecules to each of its ends. Chemical labels on the ends of these single-stranded molecules are used to attach one end to a glass surface and the other to a microscopic "handle" such as a bead. Forces can then be applied to the DNA molecule by moving the bead. The resulting extension of the DNA molecule can be measured by monitoring the position of the bead with a microscope.

Double-helix mechanics

The double helix is just that: two nucleic-acid polymer chains that are wound around one another to form a regular, right-handed helix. Each individual chain is a series of chemical units called nucleotides that are joined together by single covalent bonds. The four types of nucleotides, which contain the "bases" adenine (A), thymine (T), guanine (G) and cytosine (C), can be chained together in any order. The sugar–phosphate backbone of the structure has a directed chemical structure, which means that the base sequence along a single chain has a defined direction (see figure 2 on page 32).

However, the double helix is only stable if the two chains carry complementary base sequences because the bases only bind together in two combinations: A–T and G–C. The double helices that are found in cells (called B-DNA) obey this constraint, apart from the occasional erroneously synthesized, missing, or damaged nucleotide, and they can therefore be reconstructed from only one strand. This is the simple but remarkable principle behind DNA replication and certain types of DNA repair.

The two oppositely directed chains are bound together by hydrogen bonds and other physical interactions, which generate a binding free energy of roughly k_BT per base pair under cellular conditions, where k_B is Boltzmann's constant and T is temperature. Double helices that are more than 30 base pairs long are therefore immune to separation by thermal fluctuations, since fluctuations of more than $30k_BT$ do not occur over reasonable timescales. However, because the work that must be done to begin to separate the two strands is only a few k_BT, two strands can be taken apart gradually by enzymes without breaking the covalently bonded backbones.

At physiological temperatures, thermal motion deforms the base pairs and the backbones, causing successive bases to rock back and forth by a few degrees. The highly soluble and charged sugar–phosphate backbones tend to keep the double helix from sticking to itself in its aqueous environment, which makes long, linear DNA double helices behave like flexible polymers. In the biochemist's test tube, long double helices have no fixed shape but instead fluctuate between different random-walk conformations.

The persistence length

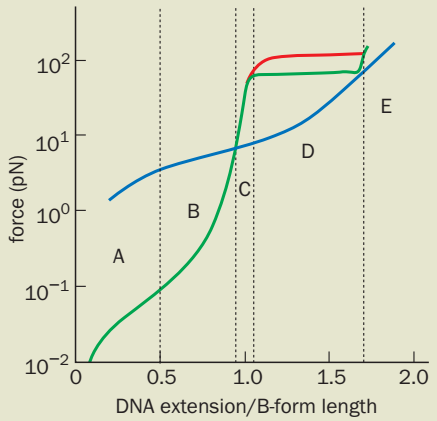
A lot can be learned about DNA by starting with simple physical ideas. The bending flexibility of the double helix, for example, can be thought of in terms of the bending elasticity of a thin rod. Classical elasticity tells us that a thin, straight rod that is bent into an arc has a bending energy $E = Bl/2R^2$, where B is the bending elastic constant of the rod, l is the length of the rod and R is the radius of the arc. Setting $R = l$ gives us the energy of a 1 radian bend along the rod, and solving for when $E \sim k_B T$ gives us the length of rod along which a thermally excited bend of 1 radian typically occurs: $l \sim B/k_B T$. This is called the persistence length, since over shorter contour lengths – lengths measured along the double helix – the double helix is essentially straight.

The persistence length of the double helix is about 50 nm or 150 base pairs. However, quoting a single persistence length is a simplification because the flexibility of DNA depends on its base sequence: G–C base-paired regions are expected to be stiffer than A–T regions because they are more strongly bound. A further complication is that some sequences are known to generate severe permanent bends along the double helix, even in the absence of thermally excited bending. Despite this, the use of a single persistence length to represent the flexibility of the double helix is a remarkably useful model.

DNA segments that are much longer than one persistence length are crumpled into random coils by thermal bending. The overall size of a random coil is much smaller than its contour length, and it can be estimated using statistical mechanics. A 50 000 base pair DNA section, which is about 17 μm long, forms a random coil with a diameter of about 1 μm. Random bending of long DNA sections is important because it brings together distant sequences on the same molecule. This allows processes such as gene transcription by distant "regulatory sequences" to occur.

The rod-bending picture suggests a way to estimate the Young's modulus of the double helix, which is defined by the ratio of stress applied to a body to the strain produced. An elastic rod with a uniform cross-section has a Young's modulus of $Y = 4B/\pi r^4$, where B is the bending modulus and r is the radius of the rod. Naively plugging in the diameter of the double helix ($r = 1$ nm) and $B/k_B T = 50$ nm, we find a Young's modulus for DNA of about 300 MPa – a value similar to that of plexiglass. Later we will describe experimental confirmation that this extrapolation of elasticity down to the nanometre scale is reasonable.

2 Applying forces to DNA



The mechanics of DNA can be studied by measuring the extent to which a single DNA molecule stretches relative to its normal B-form length when a force is applied. Double-stranded DNA (green line) requires only about 0.1 pN to be stretched to half its B-form length (A), but by applying large forces it can be stretched by more than 170 % (E). The plateau at about 60 pN (D) is due to over-stretching of the DNA. When both strands of double-stranded DNA are anchored at both ends (red line) this plateau is pushed up to about 100 pN because untwisting of the double helix is blocked. For single-stranded DNA (blue line), a force is required to extend the strand initially because it sticks to itself. The shorter persistence length of single-stranded DNA also means that it takes greater force to extend it to half its B-form length than for double-stranded DNA. Above 70 pN both double- and single-stranded DNA are extended to roughly the same length, corresponding to almost full extension of the sugar–phosphate backbones.

Experiments on single DNA molecules

In the early 1990s Steve Smith and Carlos Bustamante of the University of California at Berkeley and Laura Finzi at the University of Milan made a direct measurement of the elastic properties of single molecules of DNA. Their pioneering technique is based on powerful DNA "cut and paste" methods from molecular biology, and is now being used in different ways by many other groups.

Smith and co-workers used λ-DNA, which is the genome of a virus that infects the bacterium *E. coli*. This 48 502 base pair DNA molecule, which has a contour length of 16.3 μm, can be found in an infected cell in both circular and linear form. The latter is simply the circular form after it has been cut at a particular position along the molecule, leaving single-stranded overhangs that are 12 base pairs long (figure 1). It is thought that chromosomes – the DNA–protein complexes in cell nuclei that carry the genes and transmit hereditary information – are organized into DNA segments of about this length. This makes λ-DNA a good starting point for analysing genetic machinery.

Smith and co-workers attached a synthetic, 12 base pair single-stranded molecule to each single-stranded overhang of the λ-DNA molecule. The base sequences of these additional strands were complementary to those of the λ-DNA overhangs, allowing the two to link together or "hybridize". Chemical labels were attached to the free ends of the two synthetic molecules, which were able to stick to receptor molecules (proteins) that could be attached to arbitrary surfaces.

Using this set-up, the researchers were able to attach one end of the λ-DNA molecule to the wall of a sample cell and the other end to a paramagnetic particle with a diameter of 3 μm. This particle, or bead, provided an optically observable "handle" with which to pull the molecule. In the original experiment, a combination of magnetic fields and hydrodynamic flow was used to apply controlled forces to the bead. Today a variety of methods are used to manipulate beads that are attached to DNA molecules, such as laser tweezers and the tips of atomic force microscopes.

This simple idea of applying forces to DNA and measuring its mechanical response has led to important advances in our understanding of the mechanical properties of DNA (figure 2). Our knowledge is now so precise that the elastic response of DNA is often used to calibrate micro-manipulation experiments. However, some important questions remain unresolved, such as the exact nature of highly stretched DNA.

Stretching DNA

For very small forces below 0.1 pN, thermally excited bends in the double helix prevent the molecule from being stretched. This force scale (region A in figure 2) is determined by the ratio of the energy of one thermal fluctuation, k_BT, to the DNA persistence length, $l = B/k_BT$, where B is the bending modulus. For forces below this level, a long DNA molecule behaves like an "entropic spring", with the force going to zero for small extensions.

As the double helix is stretched from 50% to about 90% of its total B-form length, the force rises drastically from 0.1 to about 10 pN (region B in figure 2). This reflects the increasing amount of work that must be done to suppress thermal-bending fluctuations at successively shorter length scales. Calculations indicate that this mechanism should lead to a characteristic relationship between the applied force, f, and the corresponding DNA extension $f \sim (k_BT)^2/[4B(1-x/L)^2]$ as the end-to-end extension x approaches the length of the molecule L. This highly nonlinear elastic response has been observed in many experiments, and is often used by experimentalists to verify that they do indeed have a single DNA molecule in their apparatus.

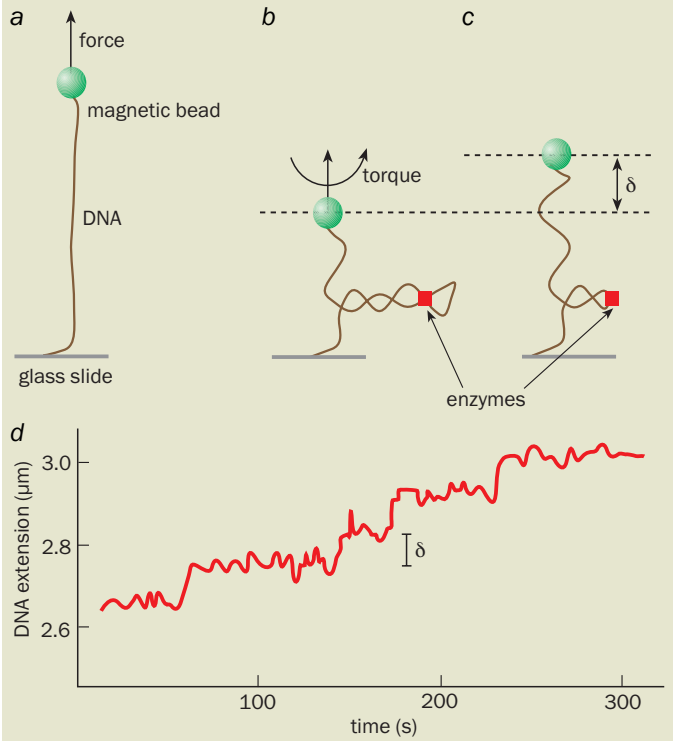
If subjected to a force of more than about 10 pN the double helix starts to deform with an extension $x \sim 1 + f/f_0$, where $f_0 \sim 1000$ pN is an elastic constant that describes the stretching of the double helix (region C in figure 2). Returning to our elastic rod picture, we expect $f_0 \sim \pi r^2 Y$, which gives $f_0 \sim 1000$ pN for a radius $r = 1$ nm and a Young's modulus $Y = 300$ MPa. In other words, the Young's modulus actually observed in DNA-stretching experiments agrees well with the value inferred from the bending persistence length.

But this linear stretching of DNA cannot continue indefinitely, and near a force of 60 pN a dramatic and sharp transition occurs (regions D and E in figure 2). The double helix changes from about 107% to nearly 170% of the length of B-DNA over just a few pN – a feature that is often used to calibrate experiments that operate in the 10–100 pN range. However, the extent of this abrupt over-stretching is a little less than what you might expect from completely straightening out the helical backbones. Successive bases in B-DNA are separated by 0.34 nm, but the distance between two bases in a fully stretched-out single strand is about 0.7 nm – over 200% of its normal length.

The structure of the over-stretched form of DNA is not yet understood. However, Philippe Cluzel, now at the University of Chicago, and Jean-Francois Leger and Didier Chatenay at LDFC-CNRS in Strasbourg demonstrated in 1998 that the 60 pN transition is suppressed when the double helix is prevented from untwisting. They also showed that over-stretched DNA is about 70% untwisted and they have suggested that the stretched form of DNA is a distinct double-helix state. On the other hand, in 2001 Mark Williams of Northeastern University, and Ioulia Rouzina and Victor Bloomfield at the University of Minnesota reported that the over-stretching transition point occurs at a lower force due to chemical conditions that destabilize the double helix. They conclude that at the end of the transition, DNA contains domains of separated "tension-melted" strands that are held together by regions of remnant base pairing.

Single-molecule methods have also been used to study single-stranded DNA (blue line in figure 2). The force–extension behaviour of single-stranded DNA is quite different from double-stranded DNA, and is complicated because it has exposed bases that tend to stick to one another at low forces. Furthermore, single-stranded DNA requires much higher tensions to be extended because it is far more flexible than the double helix, with a persistence length of roughly 1 nm (one base). Recalling that the basic force required to stretch a flexible polymer is roughly k_BT per persistence length, it follows that single-stranded DNA requires forces of a few pN to be stretched, rather than the 0.1 pN required to extend the double helix.

3 Twisting the double helix



When DNA is twisted it forms "plectonemic supercoils", just like a twisted telephone cord. Topoisomerase enzymes are able to relax these coils and this process can be studied in extension–twisting experiments. One end of a double-stranded DNA molecule is fixed while the other is twisted by attaching it to a paramagnetic particle that can be controlled by a magnetic field (*a*). DNA needs to be over-wound by about 25 turns before plectonemic supercoiling occurs (*b*). Enzymes change the topology of DNA by removing two supercoil crossings in each enzyme cycle, causing the DNA to unwind (*c*) and extend. In (*d*) a series of such jumps in extension are observed over time, which are due to individual enzyme–DNA reactions. The discrete jumps (δ) correspond to a single enzyme passing the DNA through itself via a transient break that the enzyme then re-seals.

Twisting the double helix

Just looking at the double helix, it seems obvious that it should be able to twist. Indeed, it has long been known that the double helix not only displays twisting elasticity, but also that circular DNA molecules are rather strongly unwound inside bacterial cells. Circular DNA is formed when the ends of a long, linear DNA molecule stick together, and is the form of most DNA in bacteria (including λ-DNA). Just like twisted telephone cords or garden hoses, the twisted DNA in bacteria wrap around themselves to form what are called plectonemic supercoils. This gets rid of the energetically costly DNA twisting while keeping the number of links between the two strands – a topological property – constant. In 2000 Terence Strick at Cold Spring Harbor Laboratories in New York and Jean-Francois Allemand, David Bensimon

and Vincent Croquette at the Ecole Normale Supérieure in Paris developed techniques for carrying out single-molecule DNA twisting–pulling experiments (figure 3).

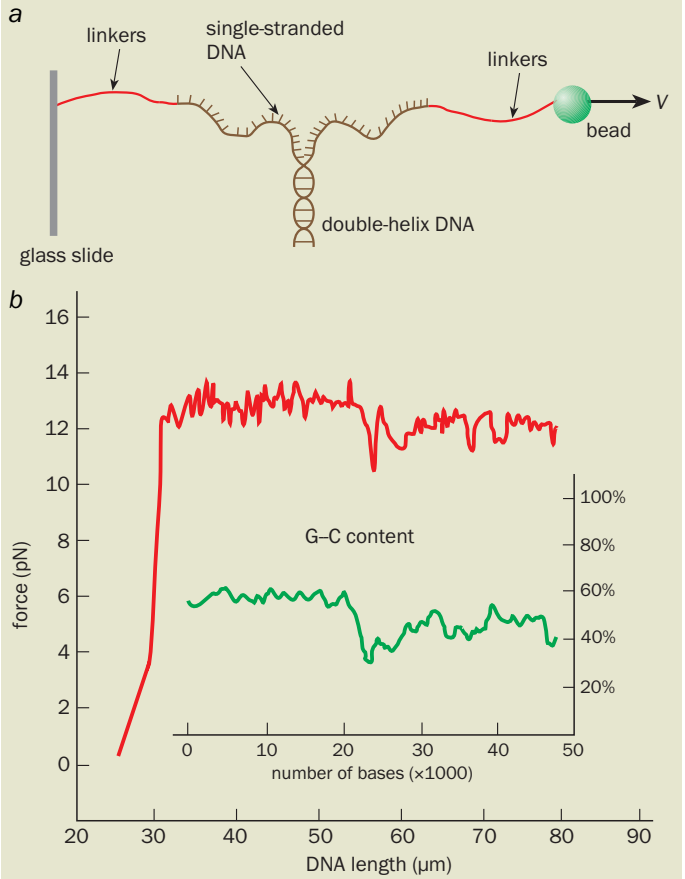
You might be wondering how circular DNA in bacteria gets unwound. The short answer is that there are enzymes that do it. An array of "topoisomerase" enzymes in cells catalyse topological changes of DNA molecules – either the exchange of single strands of DNA (type I) or the exchange of whole sections of the double helix (type II). These enzymes work their magic by first cutting the covalent bonds on the DNA backbones and then passing strands through the resulting gaps, finishing off the job by sealing the gaps. Some of these enzymes release chemical energy from the hydrolysis of adenosine triphosphate (ATP) – the chemical fuel for most sub-cellular machinery – and a few are able to transfer some of that energy into twisting their DNA hosts.

There are many mysterious properties of topoisomerases, including the apparent ability of certain type-II enzymes to recognize knots on large, circular DNA molecules. This "Boltzmann demon" property of topoisomerases was discovered by Valentin Rybenkov at the University of Oklahoma, Nick Cozzarelli at the University of California in Berkeley and Alex Vologodskii at New York University in 1997. Single-molecule experiments are now beginning to be undertaken to study the operation of these enzymes (figure 3).

The two strands of the double helix are sometimes split by enzymes inside the cell. Roughly speaking, such enzymes must apply forces in order to overcome the base-pairing interactions. Ulrich Bockelmann and Francois Heslot at the Ecole Normale Supérieure in Paris have developed a micromechanical system for unzipping the double helix, which allows a direct measurement of base-pairing interactions. This provides an excellent set-up for studying the enzymes that open up the double helix.

The force required to pull the two strands of the double helix apart is between about 10 and 15 pN, depending on the particular DNA base sequence (figure 4). The basic force scale can be understood as the ratio of the binding energy of a base pair (roughly k_BT) to the length of single-stranded DNA that is created when that base pair is opened (about 1 nm). Weakly bound A–T-rich regions only require about 10 pN to be separated, while more strongly base-paired G–C-rich regions need almost 20 pN. This was demonstrated in elegant atomic-force-microscope experiments performed in 1999 by Matthias Rief and Hermann Gaub at Ludwigs-Maximilians University in Munich.

4 Unzipping the double helix



Enzymes routinely unzip DNA molecules in biological processes. (a) Experiments can now recreate this by mechanically pulling the double helix apart at a constant velocity, *v*, causing the hydrogen bonds to be broken by the resulting force. (*b*) The force required to open the DNA molecule correlates strongly with the base sequence: A–T-rich regions with fewer hydrogen bonds require less force to unzip than G–C-rich regions.

DNA-processing enzyme machines

Given a solid knowledge of the basic mechanical properties of isolated DNA molecules, many researchers are now turning to the study of the enzymes that organize and even process the double helix. Many of these enzyme reactions depend on thermal fluctuations in the shape of DNA, and analysing the results of these experiments therefore involves a healthy dose of statistical physics.

Laura Finzi, along with Jeff Gelles at Brandeis University in the US, carried out pioneering work along these lines in 1995, studying DNA looping that is mediated by a bacterial lactose-repressor protein. This protein is able to bind to two copies of a certain DNA sequence, and when the two binding sites are on the same DNA molecule, a loop can be formed. Using this protein with a DNA molecule that could bind with it at two known sites, Finzi and Gelles were able to observe the binding and unbinding of the loop in real time. This type of interaction between two sites on the same DNA sequence occurs naturally in many gene-regulation processes, such as the repressed or enhanced reading of certain genes in the production of particular proteins. As mentioned earlier, gene expression is often controlled by distant regulatory DNA sequences, and can take place over distances of up to millions of base pairs.

Another class of DNA-processing machines are active enzymes that "walk" along the double helix. One of the most important enzymes of this type is called RNA polymerase, which makes an RNA copy of a segment of chromosomal DNA. This messenger RNA is then the blueprint for the synthesis of proteins, which are essential to all living organisms. The first micromechanical study of RNA polymerase was undertaken by Gelles and Hong Yin at Brandeis University, Michelle Wang at Cornell University, Mark Schnitzer at Lucent Technologies, Robert Landick at the University of Wisconsin and Steve Block at Stanford University. This group found that RNA polymerase is a more powerful motor than the single-molecule motors that are responsible for muscle contractions. DNA polymerase, the enzyme that duplicates DNA strands, has also been studied micromechanically, and it too turns out to be a powerful motor.

Recently a few groups have used single-molecule approaches to characterize DNA helicases – enzymes that play an important role in repairing chromosomal DNA. Based on DNA-unzipping experiments, we know that these enzymes must generate at least 20 pN of helix-splitting force to get through tough, G–C-rich DNA, and must therefore be rather powerful motors. In 2001 Piero Bianco, Steve Kowalczykowski and co-workers at the University of California in Davis visualized the separation of the two strands by a particular helicase using fluorescent dyes. The single-molecule approach allowed them to investigate the enzyme's activity by studying a sequence on the double helix that was being opened.

Another basic type of DNA processing is its packaging by a host of proteins that combine with the double helix to make it active *in vivo*. The combination of DNA and packaging protein is called a chromosome. In cells like our own, which have nuclei and other internal structures, the basic packaging unit – the nucleosome – is a combination of eight proteins with about 150 base pairs of DNA. Forces that are generated during the assembly of nucleosomes have been observed to be in the 10 pN range. The process by which strings of nucleosomes are organized into active chromosomes, however, is a major puzzle in molecular and cell biology, and micro-manipulation methods may provide important clues for solving it.

Outlook

Experiments on single molecules of DNA provide a mechanistic view of DNA-processing enzymes that is complementary to that offered by conventional solution–phase biochemical experiments. An essential feature of single-molecule experiments is that they emphasize how thermal motion and the disorder that is inherent in large bio-polymers give each enzyme–DNA interaction a unique timeframe and outcome. This basic variability at the molecular level is the origin of the complex and adaptive behaviour of living things. Over the next few years, we can expect ever more complex DNA–enzyme interactions to be observed with single-molecule methods, which will allow biologists and physicists to look even deeper into the fundamental machinery of life.

Further reading

P R Bianco *et al*. 2001 Processive translocation and DNA unwinding by individual RecBCD enzyme molecules *Nature* **409** 374–378

C Bustamante *et al*. 2000 Single-molecule studies of DNA mechanics *Curr. Opin. Struct. Biol.* **10** 279–285

S Cocco, J F Marko and R Monasson 2002 Theoretical models for single-molecule DNA and RNA experiments: from elasticity to unzipping *Comptes Rendus Physique* **3** 569–584

B Ladoux *et al*. 2000 Fast kinetics of chromatin assembly revealed by single-molecule videomicroscopy and scanning force microscopy *Proc. Natl Acad. Sci.* **97** 14 251–14 256

T R Strick, V Croquette and D Bensimon 2000 Single-molecule analysis of DNA uncoiling by a type II topoisomerase *Nature* **404** 901–904

M D Wang *et al*. 1998 Force and velocity measured for single molecules of RNA polymerase *Science* **282** 902–907

John F Marko is at the Department of Physics, University of Illinois, Chicago, IL 60607-7059, USA, e-mail jmarko@uic.edu. **Simona Cocco** is at the Centre Nationale de Recherches Scientifique, Institut de Physique, Strasbourg, France, e-mail cocco@ldfc.u-strasbg.fr