TOPICAL REVIEW

# Single-molecule studies of DNA transcription using atomic force microscopy



Phys. Biol. **9** (2012) 021001 (15pp) doi:10.1088/1478-3975/9/2/021001

TOPICAL REVIEW

# Single-molecule studies of DNA transcription using atomic force microscopy

**Daniel J Billingsley[1,2], William A Bonass[2], Neal Crampton[1], Jennifer Kirkham[2] and Neil H Thomson[1,2]**

[1] School of Physics and Astronomy, University of Leeds, Woodhouse Lane, Leeds, West Yorkshire LS2 9JT, UK
[2] Department of Oral Biology, Leeds Dental Institute, University of Leeds, Woodhouse Lane, Leeds, West Yorkshire LS2 9LU, UK

E-mail: n.h.thomson@leeds.ac.uk

Received 17 January 2012
Accepted for publication 12 March 2012
Published 3 April 2012
Online at stacks.iop.org/PhysBio/9/021001

**Abstract**
Atomic force microscopy (AFM) can detect single biomacromolecules with a high signal-to-noise ratio on atomically flat biocompatible support surfaces, such as mica. Contrast arises from the innate forces and therefore AFM does not require imaging contrast agents, leading to sample preparation that is relatively straightforward. The ability of AFM to operate in hydrated environments, including humid air and aqueous buffers, allows structure and function of biological and biomolecular systems to be retained. These traits of the AFM are ensuring that it is being increasingly used to study deoxyribonucleic acid (DNA) structure and DNA–protein interactions down to the secondary structure level. This report focuses in particular on reviewing the applications of AFM to the study of DNA transcription in reductionist single-molecule bottom-up approaches. The technique has allowed new insights into the interactions between ribonucleic acid (RNA) polymerase to be gained and enabled quantification of some aspects of the transcription process, such as promoter location, DNA wrapping and elongation. More recently, the trend is towards studying the interactions of more than one enzyme operating on a single DNA template. These methods begin to reveal the mechanics of gene expression at the single-molecule level and will enable us to gain greater understanding of how the genome is transcribed and translated into the proteome.

## Introduction

Atomic force microscopy (AFM) is an extremely versatile technique that is particularly suitable for studying biological material. Since the instrument's inception perhaps the most widely studied biomolecular sample has been deoxyribonucleic acid (DNA). AFM is one of few techniques that allows direct visualization of the molecule, and has the added advantage that sample preparation is relatively quick and simple, although there is scope for improvements in these procedures. DNA can be imaged directly on its own, without the need for complex preparation methods which can alter the molecules biological function. The instrument's versatility is reflected by the wide range of studies with regard to DNA alone that the technique has permitted. For example DNA structure, condensation and its interactions with a number of different proteins have all been studied extensively (Alessandrini and Facci 2005, Kasas *et al* 1997b, Bates and Maxwell 1993, Thomson 2006, Bai *et al* 2006). This report focuses on a single area of study, namely the study of DNA transcription by AFM. DNA transcription, the process by which DNA is 'read' or transcribed by polymerases that catalyze synthesis

 

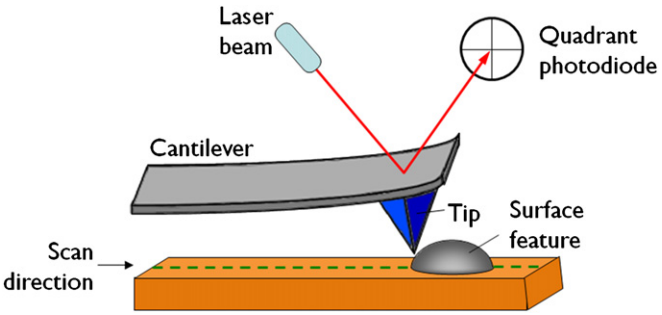


**Figure 1.** Typical set-up of an AFM. Laser light is reflected off the back of the tip onto the photodiode. The deflection of the cantilever is measured as it tracks the surface, allowing an image of the sample topography to be obtained. The tip-sample separation is controlled by feedback electronics.

of a corresponding ribonucleic acid (RNA) chain, is a vitally important step in gene expression and subsequent translation into proteins. Transcription is not a simple one-step process but is composed of a number of distinct sub-stages and it is a testament to the AFM that many of the steps in the transcription pathway can be studied. Being a single-molecule technique, AFM allows new insights into transcription to be gathered, results that are not possible with some of the more traditional ensemble biochemical techniques. This review starts with a brief overview of the instrumentation and then describes typical sample preparation procedures for studying DNA on mica. It focuses on some of the more practical aspects of studying DNA by AFM, before reviewing the body of work on DNA transcription. Finally, we discuss future developments which should allow greater insights into transcription to be obtained by use of AFM.

## Scanning probe microscopy

The atomic force microscope (AFM) belongs to a family of instruments known as scanning probe microscopes (SPM). These devices all consist of a probe, specifically designed to measure a particular physical quantity. Images are obtained by mechanically moving the probe with respect to the sample surface. Probe–surface interactions are recorded at discrete points, allowing a digital image of the surface to be constructed The main advantages over other techniques are the ability to investigate biological samples in their native states, without the need for complex sample preparations (e.g. fluorescent or heavy metal staining), that may damage or hinder the samples' biological activity. It is possible to perform imaging in air, vacuum and also liquid (Hansma *et al* 1994, Putman *et al* 1994), with the latter allowing measurements to be taken in an environment compatible with physiological conditions.

### *The atomic force microscope*

Central to AFM is the force measuring probe, which consists of a cantilever with an integrated tip at its end that interacts with the sample surface (see figure 1). These cantilevers are micro-fabricated and vary in design depending on the environment and imaging mode used. AFM works by observing deflections of the cantilever in the nanometre range caused by surface

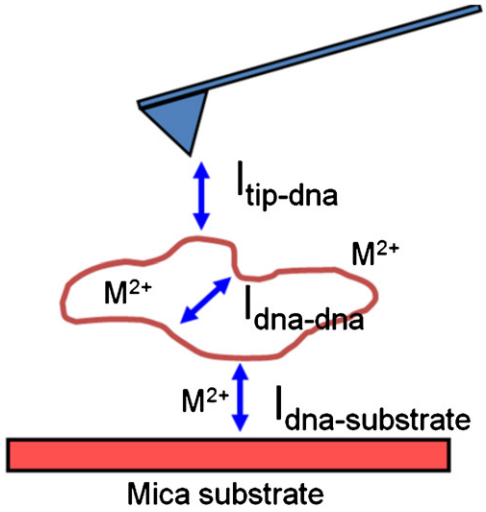


**Figure 2.** Interactions ($I$) involved with imaging DNA with AFM. For effective imaging the interaction between DNA and the substrate must be greater than the interaction of the molecule with the tip (i.e. $I_{\text{dna-surface}} > I_{\text{tip-dna}}$, to prevent the molecule being disturbed by the motion of the tip. In order to image DNA effectively divalent cations ($M^{2+}$) are often introduced into the deposition buffer, with the view to reducing the natural repulsion between negatively charged DNA and mica.

forces of nN or below. To detect small changes in the cantilever deflection or amplitude it is important to have a sufficiently sensitive detection method, the most common of which makes use of an optical lever (Meyer and Amer 1988). The movements of the tip are measured using a quadrant photodiode. In the first instance, a laser is focused on the cantilever and the reflected beam hits the photodiode in its centre when the cantilever is in free space. Any deflection of the cantilever, when the tip is in the presence of the surface, causes a shift of the laser spot across the interfaces of the split photodiode. These power differences can be used to detect both bending and torsional motions of the cantilever. To determine the force between the tip and sample one needs to accurately measure the cantilever stiffness in any mode and the optical lever sensitivity. As the intermolecular forces that cause the cantilever deflection, and give rise to imaging contrast, are very short ranged, the cantilever must be positioned vertically in very close proximity to the sample surface. In addition to this, very precise movements in the $X$ and $Y$ directions are necessary during scanning. Most AFM systems achieve such positional accuracy by using piezoceramic transducers, which allow the movements of the probe with high precision down to the sub-nanometre range.

The original method of AFM operation is known as contact mode; in which the tip remains in close contact with the sample during its scan. Early AFM studies of DNA utilized contact mode (Hansma *et al* 1991, Lindsay *et al* 1989); however, it took some time before routine imaging of the molecule could be performed in ambient conditions (Bustamante *et al* 1992). Problems were caused by a number of technical issues, including (1) the substrates or support surfaces used, (2) the standard of tips available at the time and (3) the mode of imaging. These issues are, to a greater or lesser extent, inter-related, essentially as a consequence of Newton's

third law, because the AFM tip typically exerts forces on the sample at the same time as detecting a force-based response from it. For successful imaging of single biomolecules, the interaction of the molecule with the supporting surface needs to be stronger than that between the AFM tip and molecule (figure 2).

Commercial production of AFM tips is still being improved and optimized; however, tips with a larger radius will interact with the sample over a wider area, which can lead to greater disturbance of soft matter and weakly bound molecules. This is particularly relevant to contact mode imaging, where the continuous tracking of the tip across the surface leads to the exertion of large shear forces, which can displace weakly bound particles or damage soft samples such as DNA. This problem is typically made worse by performing imaging in air, where the capillary neck provides a large contribution to the shearing force. This issue can be reduced by using sharper tips to minimize the dimensions of air–liquid phase (Bustamante *et al* 1992). It has been shown that it is possible to image DNA repeatedly using contact-mode, as long as the humidity is kept below a critical value (Thundat *et al* 1992), such that the magnitude of an attractive capillary force, sucking the AFM tip down onto the surface, is avoided. Reliable imaging of DNA by contact mode was not possible with humidity above 40% RH (Thundat *et al* 1992, Vesenka *et al* 1992), which correlates with the humidity where water condenses to form thick films on the surface (Hu *et al* 1995). It is also possible to operate the instrument in bulk aqueous fluid allowing the imaging to be performed in a state more akin to physiological conditions, whilst also allowing dynamic processes to be investigated (Hansma *et al* 1994, Putman *et al* 1994).

A significant advance came with the development of imaging modes where the cantilever was vibrated, typically close to its resonant frequency, in a range of methods known as dynamic AFM (dAFM) (Garcia and Perez 2002). The contact of the tip to the surface is intermittent and in the normal direction to the surface plane, almost completely eradicating the damaging shearing interactions. In dAFM, it is easier to control the cantilever to operate in non-contact imaging above the surface than contact mode, where instabilities are prevalent. Typically, for standard imaging of biomolecular and nano samples the AFM is operated such that the tip makes periodic mechanical contact with the surface, in modes colloquially known as either intermittent-contact or tapping-mode AFM (Zhong *et al* 1993). Intermittent contact with the surface at the end of each oscillation leads to the dissipation of energy from the system and a decrease in oscillation amplitude. The amplitude can be used as a set-point parameter, so that feedback electronics can act to adjust the height so as to maintain the set-point at its operator pre-set value. The amount the amplitude must be modulated can be used to gather topographic information of the surface, thus this form of imaging control in dAFM is known as amplitude modulation AFM (AM AFM).

The contact force the tip exerts at the end of each cycle in dAFM can often be considerably smaller than in contact mode and is typically of the order of nN (Zhong *et al* 1993). The smaller forces exerted and lack of shearing mean that tip-induced sample deformation is minimized and soft samples can be imaged reproducibly. The use of AM AFM for DNA imaging, most usually in tapping mode, has almost entirely replaced contact mode. The work on DNA transcription reviewed in this article has been exclusively acquired with tapping-mode imaging, more robustly known as repulsive force regime AM AFM, in either ambient air or aqueous liquid environment.

## Immobilization of DNA molecules on mica surfaces

By its very nature AFM is a surface scanning technique and as such any molecule that is to be investigated must be deposited on to a suitable substrate prior to imaging. There are therefore inevitably a number of criteria that the support must satisfy for it to be suitable. The substrate in question must have a low surface roughness, such that molecules of interest can easily be distinguished from any spurious features of the support surface. It is crucial that there is a strong interaction between the surface and molecule, so that the sample will move from being in solution to being adsorbed onto the substrate. Strong binding to the substrate surface is also required so that the sample will resist the motion of the moving tip, without being damaged or detached from the surface. To study dynamic processes in liquid it may often be necessary to modulate the sample–surface interaction by ion exchange, moving from a weak interaction to allow molecular movement to a stronger interaction for effective imaging (Thomson *et al* 1996).

Mica has been found to be a reliable and convenient substrate for imaging of DNA, and is now the most commonly used support surface (Vesenka *et al* 1992). Most biomolecules, such as DNA, are hydrophilic in nature and thus a hydrophilic material such as mica provides an ideal support. Mica is the collective name for a number of different silicate-based minerals. Muscovite mica is used for AFM studies and has the structure $KAl_2(AlSi_3O_{10})(OH)_2$. The main characteristic of mica materials is the perfect cleavage of the basal layers. This is ideal for AFM studies where the upper layer can be removed by adhesive tape to provide a clean surface, free from impurities each time an experiment is to be performed.

Preparing a DNA sample for AFM imaging typically follows a three-step process: first a small amount of DNA in buffer is deposited onto a mica disc, before being left to for a period of time to bind. Finally, after incubation, the sample is rinsed with a small amount of pure water and dried with a weak flux of argon or nitrogen gas. The deposition of DNA onto the substrate is a two-step process; beginning with the transport of molecules to the surface and ending with binding to the mica surface (Lang and Coates 1968, Rivetti *et al* 1996).

The most straightforward method of sample preparation uses physisorption processes, where the environment and the surface are controlled to immobilize the molecule via the natural forces present. Surface binding can also be achieved through a chemical strategy where the molecule is covalently linked to the surface. This can be achieved by chemically functionalizing the substrate with positively charged groups, such as amines ($NH_2$). Aminosilanes,

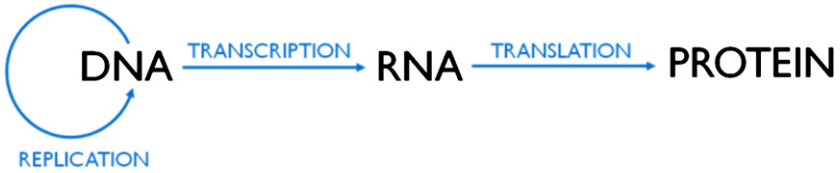


**Figure 3.** The central dogma describes the residue by residue transfer of sequence information between the information-carrying biopolymers. This ultimately leads to the expression of genes into proteins.

particularly aminopropyltriethoxysilane (APTES), have been used to create positively charged surfaces, meaning that negatively charged DNA will interact strongly with the substrate (Crampton *et al* 2005, Lyubchenko *et al* 1993, Shlyakhtenko *et al* 1998, 1999). The advantages of using APTES-mica is the strength of the interaction between itself and DNA, meaning it is possible to stably image the molecule in both air and under aqueous buffer. The strength of binding is also not dependent on the presence of cations, meaning constraints on buffer and the system studied are negated. These surfaces, however, do lead to kinetically trapped molecules in which DNA cross-overs are allowed and therefore may not be so suitable for studying translocation events such as those in transcription (Pastre *et al* 2003, Rivetti *et al* 1996). The majority of the work on AFM of transcription has been carried out on mica surfaces modified through ion exchange to facilitate DNA binding (Crampton *et al* 2006a, Rivetti *et al* 1999).

## Introduction to transcription and RNA polymerase

Transcription is one of the key steps in the process of gene expression and its study by AFM is the focal part of this report. DNA's main function is in the storage of genetic information used for encoding proteins, and the ability to pass-on this information to future generations. All cellular processes depend on proteins in some way, and as they carry such an importance for life they must be constantly produced in a regulated and reproducible manner. The information contained within a gene is passed on during the process of transcription; a process whereby RNA polymerases (RNAPs) use DNA as a template for the synthesis of an RNA chain (see figure 3). The RNA is a direct copy of the coding DNA strand and includes exons and introns. The newly synthesized RNA then undergoes post-transcriptional modification involving removal of the 5′ terminal and capping by a methylated guanosine, cleavage of the 3′ end and polyadenylation, and splicing where introns are removed, and the exons are connected to form a messenger RNA (mRNA) chain. The mature mRNA then moves outside the cell nucleus where it is translated into chains of amino acids (proteins) by the ribosomes.

DNA transcription is carried out by the enzyme RNAP, and was first described in 1960 (Stevens 1960, Hurwitz *et al* 1960, Hurwitz 2005). RNAPs can occur as both single- and multiple subunit enzymes. Bacteriophages and mitochondria contain RNAPs of the single-subunit variety, whilst polymerases from bacterial, archaeal and eukaryotic cells are representative of the multi-subunit class. Of the multi-subunit RNAPs, perhaps that of *Escherichia coli* is the best-studied. These enzymes are constructed from four sub-units: $\beta$ containing 1407 amino acids, $\beta'$ (1342 amino acids), and a dimer of $\alpha$ (329 amino acids) (Darst *et al* 1989).

Zhang *et al* used x-ray crystallography to reveal the structure of RNAP from *Thermus aquaticus* (Taq) a thermally stable type, with a resolution of 3.3 Å (Zhang *et al* 1999). The molecule is characterized by its 'crab-claw' shape, housing a groove running the entire length of the inner surface. This is a distance of approximately 30 Å, allowing around nine base pairs (bp) of DNA to be accommodated. The active site of the enzyme is located deep within the molecule, where there are also binding sites for catalytic Mg(II) ions. In addition to the main channel there is a secondary pore, which allows individual nucleotide triphosphate (NTP) molecules access to the active centre of the RNAP. The crystal structure of *Taq* RNAP corresponds well with the structure of *E. coli* RNAP obtained by electron crystallography (Polyakov *et al* 1995), implying that functional data obtained from experiments on *E. coli* RNAPs might be rationalized using the high resolution structure of *Taq* polymerase. A similar morphology is also shared by the structure of RNAP II as seen in figure 4.

The process of transcription can be separated into a number of main stages (summarized in figure 5):

- Initiation
- Elongation
- Termination

In the initiation stage, the RNA polymerase is able to recognize a specific DNA sequence, termed the promoter region. This region lies just upstream of the main coding sequence of the gene. At this point the polymerase acts to melt, or separate, the two DNA strands to form a 'transcription bubble'. Certain DNA sequences are highly conserved and are found in the promoter region of most genes. For example, the TATA box is an AT-rich sequence, occurring around 30 bp upstream from the transcription start site (Corden *et al* 1980). This sequence is recognized by specific proteins called $\sigma$-factors. A major protein involved in the *E. coli* transcription cycle is $\sigma^{70}$, with the superscript denoting the molecular weight in kDa (Ring *et al* 1996, McClure 1985). DNA binding-proteins such as transcription factors are typically a prerequisite for the cell to become transcriptionally active.

Once the RNAP, either alone, or in conjunction with a $\sigma$-factor, becomes bound to its promoter region, the complex undergoes a conformational change. The RNAP acts to orientate itself so as to trap the DNA into its 'claw'. During this process the system moves from what is termed a closed promoter complex (CPC), to an open promoter complex (OPC), with the nomenclature denoting the state of the 'transcription bubble'. The initial binding is to double-stranded or closed DNA but the reorientation of the RNAP is

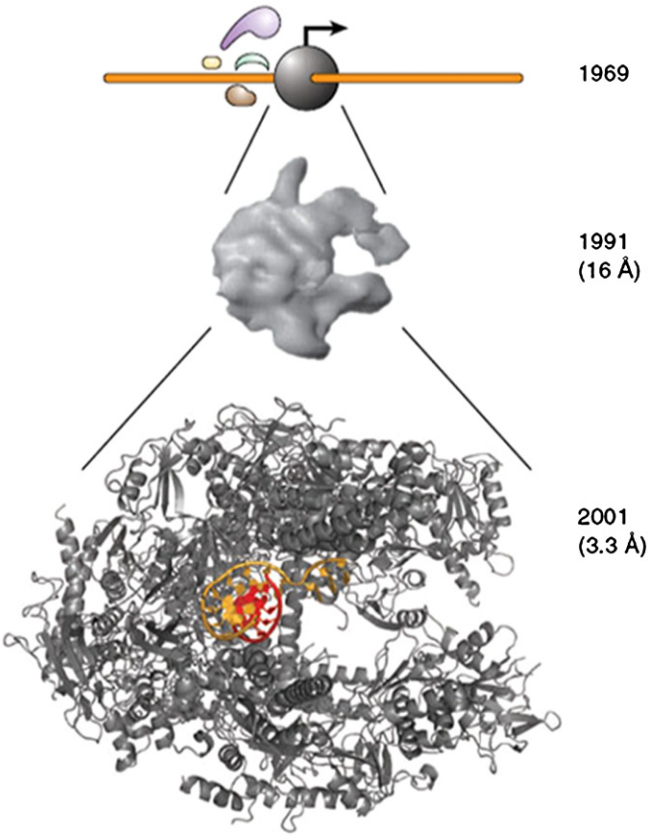


**Figure 4.** Development of knowledge in the structure of RNA polymerase II through time using molecular biology, electron and x-ray diffraction (Ansari 2007). It was first isolated and characterized as the enzyme that transcribes genes in 1969. Transcription factors (structures to the left of the RNAP) mediate the recruitment of RNAP onto the chain to initiate transcription. X-ray crystallography has allowed the interior of the enzyme to be probed to 3.3 Å resolution, and is characterized by its 'crab-claw' shape. The enzyme catalyzes the unwinding of the DNA helix during transcription. The structure resembles that of a hand, with a region of single-stranded DNA lying in its 'palm'. NTPs can enter through a channel in the interior of the enzyme. This are joined together using DNA as a template to form a complementary chain of RNA, which exits the enzyme under its 'thumb' which is used to clamp across the double-stranded DNA helix. Reprinted by permission from Macmillan Publishers Ltd: [Nature Chemical Biology] (Ansari, A.Z., Chemical crosshairs on the central dogma), copyright 2007.

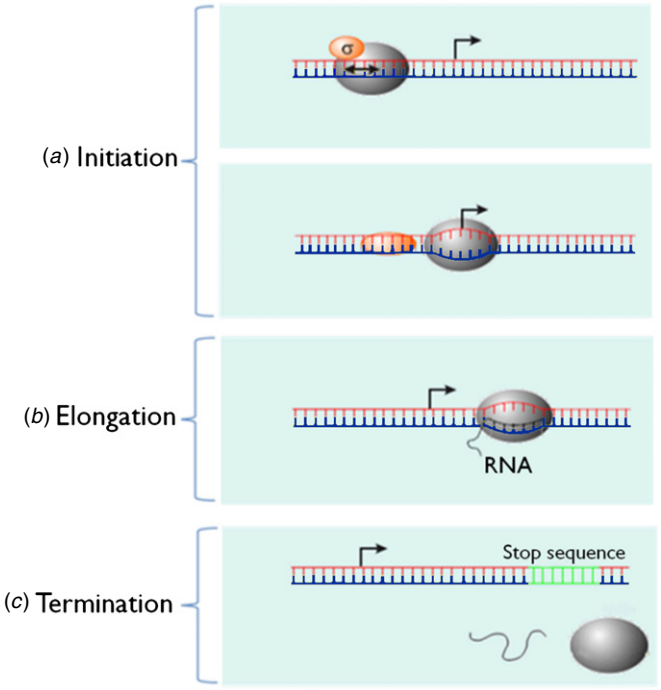


**Figure 5.** The transcription cycle of *E. coli* RNAP. The RNAP is shown as the grey sphere, $\sigma$-factor as an orange sphere, and the two DNA chains are indicated in red and blue. (*a*) Initiation: RNAP recognizes and forms stable contacts with a sequence of DNA known as the promoter, recognition is aided by additional protein regulators known as $\sigma$-factors. The promoter is located by sliding along the DNA, before this sequence is melted in the binding region to form a transcription bubble or open promoter complex. (*b*) Elongation: the RNAP moves away from the promoter and loses its $\sigma$-factor, and begins to synthesize RNA using single-stranded DNA as a template. (*c*) Termination: finally, the RNAP encounters a stop sequence, which destabilizes the DNA–RNAP complex leading to transcription termination, and release of the RNA molecule.

accompanied simultaneously with the unwinding of a 10–15 bp segment of DNA, to form the 'transcription bubble' or open complex (deHaseth *et al* 1998).

Next the complex undergoes a process known as abortive initiation, whereby competition exists between short RNA oligo release and continuous NTP incorporation ultimately leading to the completed messenger RNA chain (Carpousis and Gralla 1980). The $\sigma$-factor remains bound to impart stability to the RNAP and prevent disengagement from the template. Once the DNA transcript reaches ~12 nucleotides, the RNA–DNA complex becomes stable and the RNA loses contact with its $\sigma$-factor, escaping the promoter. This point marks the end of the initiation stage and the beginning of chain elongation (Hsu 2002).

During elongation the RNAP proceeds along the DNA template, synthesizing a chain of RNA as it goes. A ternary complex consisting of RNAP, DNA and RNA is formed. The RNA chain is constructed by a polymerization reaction, in which nucleotide triphosphate (NTP) monomers are hydrolyzed, supplying the energy to form a long chain. This RNA is a direct copy of the coding DNA strand, with the notable exception that the base thymine in DNA is replaced by uracil (*U*).

During termination, transcription is halted, the RNAP dissociates from the DNA template and the RNA transcript is released. This can be caused by the RNAP encountering a particular DNA sequence, known as a terminator. This situation is called intrinsic termination, in which a sequence encodes an RNA that forms a hairpin structure, which acts to destabilize the elongation complex. Termination can also be controlled by protein factors (Das 1993). Whilst still in the nucleus, the RNA undergoes post-transcriptional processing, including capping, polyadenylation and splicing, before translation of the RNA transcript occurs outside the nucleus, which ultimately leads to the construction of a specific protein.

## Practical aspects of studying transcription with AFM

As a high resolution microscope, AFM can be used as a single-molecule technique and offers advantages over more traditional biochemical techniques for the study of DNA transcription, which have typically been performed on large populations of molecules, where measurements of properties represent ensemble averages. This can lead to a situation where events occurring at a single genetic locus are obscured due to lack of synchronicity between different molecules in the system. With AFM it is possible to observe each member of the population individually, and to build up an overall distribution of results. While this is presently a time-consuming process, AFM allows molecular structures to be observed directly, and can provide new information on the spatial arrangements of DNA, RNA and protein during the various stages of the transcription cycle.

Studies of transcription via AFM typically follow very similar sample preparations (Kasas 2003). This usually involves the completion of the reaction in a test-tube, in which the RNAP, specially designed DNA templates containing promoter sequences, and three of the four nucleotide triphosphates (NTPs), are mixed together. The RNAP associates with the promoter region, melting the target sequence and forming an open promoter complex. The presence of the NTPs allows transcription to be performed *in vitro* and the OPC will go on to form an elongation complex. A chain of nascent RNA will begin to be constructed out of NTPs. This period of elongation continues until the RNAP encounters a base on the template DNA which codes for the NTP that is missing from the solution. This point is known as the stall site and the structure consisting of the protein, small chain of RNA and template DNA, is known as a stalled elongation complex (SEC).

Static studies of general structure can be performed readily in ambient conditions. Divalent cations, which help to facilitate DNA binding, are added to the buffer containing the DNA, before a small amount is transferred onto mica and allowed to incubate for a period of time. The surface is then rinsed, and dried in a weak flux of gas. This method provides a 'snap-shot' of the complex and allows the further exploration of OPC and SEC structure (Rivetti *et al* 1999, 2003, Rees *et al* 1993).

The conformation that DNA adopts on the mica surface is a direct consequence of the various forces present during the binding process. Rivetti was able to experimentally show that there existed two extremes of surface-bound DNA conformation, where the molecules can said to be either surface equilibrated or kinetically trapped (Rivetti *et al* 1996). During surface equilibration the molecules are bound to the surface but once touching the surface they are free to diffuse laterally and take up a minimum energy conformation. These conformations are mediated by long-range, weak interactions, whereas short-ranged and stronger interactions will cause the molecule to be kinetically trapped. During kinetic-trapping, if any part of the chain touches the surface it is irreversibly pinned at that location on the surface and the rest of the molecule collapses onto that point. It is no longer possible for the molecule to move sideways and as such, the conformation describes a 2D projection of the DNA's 3D structure. For example, it has been shown that the use of Mg(II) in the deposition buffer leads to the surface equilibration of the molecules, whilst Ni(II) pre-treatment of the mica surface causes the DNA to take up a kinetically trapped conformation (Billingsley *et al* 2010, Pastre *et al* 2006, Rivetti *et al* 1996).

Dynamic studies of transcriptional elongation can also be performed using AFM but these measurements require the AFM to be operated under liquid. Imaging the process of transcription directly presents a paradox that must be overcome. Under bulk liquid, molecules are mobile and interacting dynamically with the support surface. Lateral frictional forces between the molecule and surface are reduced, leading to a situation where molecules can diffuse laterally across the surface and even detach entirely. In addition to this, the tracking of the tip over the surface takes a relatively long time. If these diffusional motions are quicker than the time resolution of the instrument, then the tip becomes unable to resolve the molecule properly and the image obtained will be blurred. However, a strong anchoring of the molecule to the substrate is unfavourable for the study of dynamical processes such as transcription. For the RNAP to properly transcribe the template, the DNA must be allowed some freedom to move and rotate.

This problem can be circumvented by use of a gravity driven flow-through system, permitting the exchange of buffers in the liquid cell of the AFM. Using this method it is possible to achieve reversible binding of DNA to the mica surface, essentially by controlling the concentration of divalent cations, in the environment around the DNA (Thomson *et al* 1996) (see figure 6). The apparatus used consists of containers suspended above the AFM, containing regulators to adjust the flow rate. Buffer solution is constantly being driven into the system by the influence of gravity and the regulator is used to switch from one solution to another. In the study of transcription, two solutions are used: an 'imaging' buffer and a 'transcription' buffer. The former allows enough adhesion for complexes to be imaged properly, whereas the latter contains the NTPs and also promotes detachment, allowing the RNAP to move with respect to the DNA and transcribe it.

## AFM studies of transcription

### *Promoter search*

For transcription initiation the polymerase must associate with the promoter region, a relatively short recognition sequence that delineates the start of the gene. If the polymerase were to diffuse in three dimensions until it encountered the promoter site, then the time taken would be prohibitively long and transcription would not proceed at the required rates for adequate protein expression. It has been suggested that if a diffusion-controlled promoter search were to occur in a space of reduced dimensionality, then the efficiency could be increased by orders of magnitude (von Hippel and Berg 1989).

A number of mechanisms have been suggested to explain the increased binding rates associated with promoter regions.

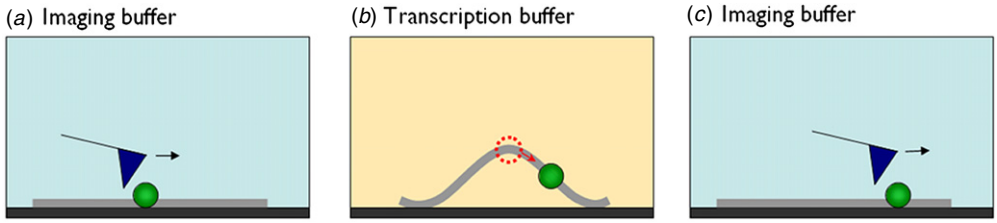


**Figure 6.** Method for studying transcription dynamically in liquid with AFM. DNA is shown as a grey line, and RNAP as a green sphere. (*a*) Use of an imaging buffer allows effective binding of DNA to the substrate, and leads to reliable imaging. (*b*) Substitution with a transcription buffer enables the DNA to detach from the surface and RNAP translocation. (*c*) Subsequent imaging in the original buffer shows that the RNAP has moved along the chain. This method produces time-lapse images of transcription.

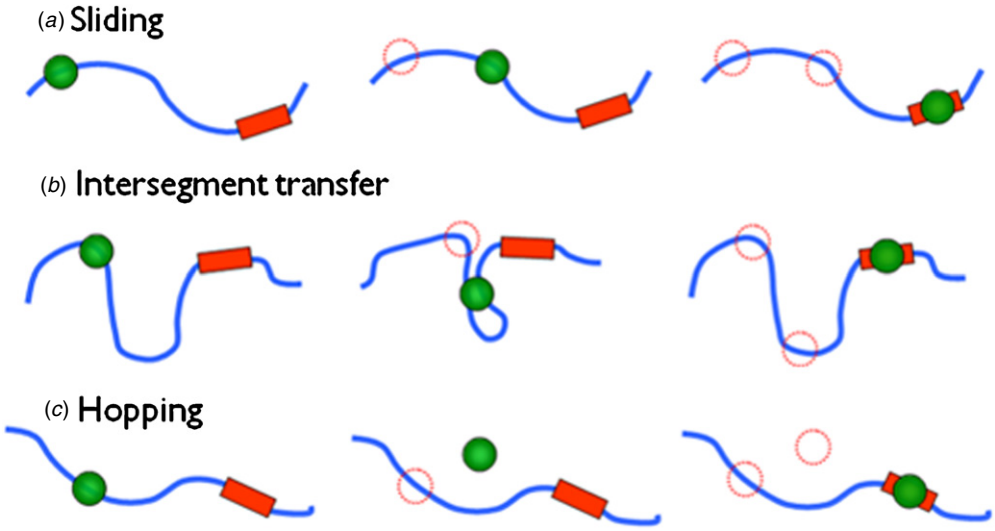


**Figure 7.** Schematic of the different methods of promoter location observed by AFM (Guthold *et al* 1999, Bustamante *et al* 1999). The DNA is shown as a blue line, the RNAP as a green sphere, and the promoter region as a red rectangle. (*a*) In sliding the RNAP diffuses one-dimensionally along the DNA contour. (*b*) In intersegment transfer the RNAP is able to reach a distant site on the DNA via the formation of an intermediate complex. (*c*) During hopping the RNAP loses contact with the DNA, before re-attaching further along the chain.

If the RNAP were to have an affinity for non-promoter DNA, then the likelihood of non-specific binding to the DNA would be high. The polymerase could then undergo one-dimensional diffusion along the chain, until it reached the target sequence at the promoter (Berg *et al* 1981). This method is known as *sliding*. Another mechanism, known as *intersegment transfer*, involves the transfer of the polymerase from one position on the template to a more distant segment by means of an intermediate state where the DNA is looped, such that the protein is associated both to the initial and distant regions at the same time (von Hippel and Berg 1989). Multiple transfer events occur until the promoter region is reached. A final process, called *hopping* has also been put forward, in which the polymerase bounces along the template DNA until it finds the target site (von Hippel and Berg 1989). All three of these mechanisms serve to reduce the dimensionality of the promoter search process.

Guthold *et al* have been able to observe the one-dimensional diffusion of *E. coli* RNAPs along a promoterless DNA fragment by AFM (Guthold *et al* 1999). This experiment was performed in liquid, in conditions that could be controlled to allow both diffusion and imaging to occur interchangeably, in this case by varying Mg(II) concentration. The protein was observed strongly bound to the surface but the DNA regained a certain degree of mobility and could diffuse laterally. Sequential images of these complexes were taken, in which the DNA was seen to slide back and forth beneath the enzyme. The position of the protein on the chain in successive images was measured. The mean diffusion distance, obtained from measurements on a number of molecules, was found to be approximately proportional to the square root of time, typical of diffusion-controlled processes.

In conjunction with sliding, AFM studies of non-specific complexes have also observed intersegment transfer and hopping (Bustamante *et al* 1999). However, these events are not nearly as common as sliding. Images showed an intermediate structure consisting of an RNAP molecule bound to two segments of the same DNA molecule. In conjunction with this ternary complex, the intersegment transfer event is characterized by the formation of a tight hairpin structure. The polymerase was then transferred to another segment, roughly 100 bp away from its initial position. Other images were collected where an initially DNA-bound polymerase was observed to dissociate from the DNA, before re-associating at a point further along the chain (Bustamante *et al* 1999). These 'hopping' events could occur repeatedly until the target site is located. These events were preceded and followed by sliding, so as such, it is possible that a combination of sliding,

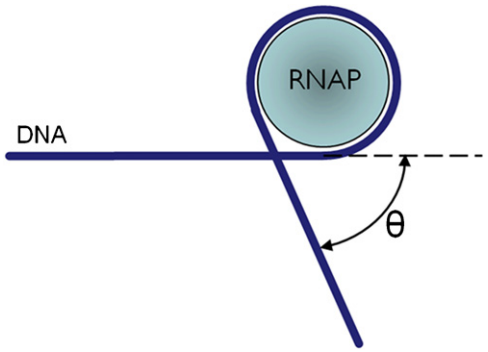


**Figure 8.** Model demonstrating how DNA can wrap around an RNAP core, and how the characteristic bend angle is defined.

intersegment transfer and hopping are responsible for the most efficient promoter search (see figure 7).

However, as a cautionary note, the lifetime of the non-specific complexes studied in AFM studies is much longer than bulk studies performed in solution (Herbert *et al* 2008, Singer and Wu 1987). The fact that the DNA–protein complex is surface-bound in AFM studies would account for this discrepancy. The mica surface imposes a conformational constraint on the complex, slowing the reaction and preventing the enzyme from dissociating. On the other hand, it can be argued that as living systems are filled with surfaces (e.g. membranes), AFM studies of molecules bound to surfaces could be more biologically relevant than performing reactions in test tubes (Hansma 2001).

### *Open promoter formation*

Promoter recognition and binding by the polymerase signifies the beginning of transcription. The initial stage of the process is typified by conformational changes occurring in the polymerase (Durniak *et al* 2008), and the unwinding of a region of DNA at the promoter to form a transcription bubble. The resulting complex is relatively stable, and known as the open promoter complex (OPC). Using AFM it is possible to study the structure of the OPC, and contrast it to the initial closed state that the protein and DNA exists in.

The specific binding of the RNAP at the promoter site alone is not always enough for the formation of open complexes. Other DNA-binding proteins ($\sigma$-factors) help to enhance the reaction. A major $\sigma$-factor in *E. coli* is $\sigma^{70}$, which can act solely to unwind the template without the need for any additional protein factors (Dame *et al* 2003). The *E. coli* RNAP$\cdot\sigma^{70}$ holoenzyme provides a relatively simple system with which to study the structure of OPCs with AFM.

Observations of RNAP$\cdot\sigma^{70}$ OPCs with AFM have shown that the DNA template always appears bent in the region bound to the polymerase (Rees *et al* 1993, Rivetti *et al* 1999) (see figure 8 for an explanation). No significant curvature is observed in DNA alone over sufficiently long length scales (Wiggins *et al* 2006) and it can be concluded that these bends are caused by the conformational changes associated with the open complex formation. AFM can quantify the bend angle projected into 2D, for example, Rees *et al* examined complexes formed at the $\lambda P_L$ promoter and obtained a mean bend angle of 54° (Rees *et al* 1993), while Rivetti *et al* looked at the $\lambda P_R$ promoter and observed a wide distribution of bend angles, with the DNA being bent between 55° and 88° (Rivetti *et al* 1999).

Rivetti *et al* also performed a detailed analysis of the contour lengths of many DNA molecules and found that in the case where RNAPs were bound to the promoter in an open conformation, the length of the fragment was reduced by ∼30 nm (∼90 bp) relative to free DNA. This, together with the bend angle measurements, enabled a model of the open complex to be put forward. The reduction of contour length is consistent with a model in which the promoter DNA is wrapped around the polymerase by 300°. For transcription to initiate, 12 bp around the transcription start site must be unwound to form the transcription bubble. Rivetti suggests that the super-helical left-handed twist, provided by the wrapping of the promoter region around the protein core, could provide for the conversion of negative writhe into local untwisting, aiding promoter clearance (Rivetti *et al* 1999).

AFM studies have also been completed in which transcription-factor $\sigma^{54}$ was used instead of $\sigma^{70}$ during transcription initiation (Schulz *et al* 1998, Rippe *et al* 1997). This method is slightly more complicated, as the RNAP$\cdot\sigma^{54}$ requires additional activators (e.g. nitrogen regulatory protein C (NtrC)) to unwind the DNA to initiate transcription. Upon ATP hydrolysis, the activator protein catalyzes the transition of the holoenzyme from a closed complex to an open complex. Rippe *et al* were able to carefully control the reaction occurring in a test-tube, such that when the sample was imaged under AFM, different stages in the transcription activation pathway could be visualized (Rippe *et al* 1997). Both closed and open complexes were imaged, along with an intermediate stage. The intermediate complex consisted of contacts between RNAP$\cdot\sigma^{54}$ at the promoter and the NtrC, with the intervening DNA forming a hairpin loops. Looping provides a way for the distal activator protein to access the promoter region containing its associated complex, with the resulting interaction allowing transcription activation to take place. As both closed and open complexes were imaged, direct comparisons can be made between the two. Both kinds of complexes caused the bending of DNA in the region around the promoter, with closed complexes exhibiting an apparent bend angle of 49° ± 24°, and open complexes yielding an angle of 114° ± 18° (Rippe *et al* 1997). These differences were interpreted by explaining that the transition from a closed complex to an open complex during transcription activation was accompanied by structural changes in the complex.

### *Elongation studies*

The elongation stage begins when the RNA transcript reaches around 12 nucleotides in length. After a stage of abortive initiation, the polymerase breaks free of the promoter region and any associated $\sigma$-factors, forming a stable elongation complex that is strongly bound to the DNA template (Hsu 2002). The RNAP moves along in a processive manner, incorporating NTPs into an ever growing nascent RNA chain. If transcription is initiated with only three of the four types of

NTPs present in solution then the enzyme will travel along the template until the first occurrence of the missing base, with the resulting complex being a stalled elongation complex (SEC).

Rees *et al* performed studies on both OPCs and SECs, with measurements of the protein-induced bends on the template providing a comparison of the two different structures (Rees *et al* 1993). The mean bend angle in elongating complexes was found to be 92°, compared to 54° in open complexes. These differences were attributed to structural changes in the RNAP upon its maturation from an open to an elongating complex: a hypothesis supported from studies by Rivetti *et al* (2003). SECs of *E. coli* polymerase and template DNA were formed and transferred onto mica. Measurements of the DNA contour length were undertaken, and it was found that the template exhibited a ∼22 nm reduction in length (Rivetti *et al* 2003). This was lower than earlier measurements performed upon open complexes, whereby a contour length reduction of about 30 nm was observed (Rivetti *et al* 1999). The reduced wrapping was attributed to the loss of contacts between the DNA and polymerase during the transition from initiation to elongation (Rivetti *et al* 1999). Sequence-specific contacts with the promoter region are essential for initiation, and are the main contribution to wrapping. It is believed that some of these interactions are lost upon promoter escape (Dame *et al* 2003).

More dynamic aspects of transcription elongation can be followed by performing the imaging under aqueous fluid, in the manner described previously in this report. Kasas *et al* performed tapping-mode AFM on a DNA fragment containing the $\lambda P_R$ promoter site (Kasas *et al* 1997a). SECs were formed in solution and subsequently deposited onto mica. A transcription buffer containing all four of the NTPs needed for elongation was injected into the fluid cell. The concentration of each NTP was kept deliberately low (0.5–5 $\mu$M), in order to circumvent one of the inherent drawbacks of AFM. Data acquisition in AFM, as with other scanning probe microscopy techniques, is relatively slow with an image typically taking tens of seconds to be obtained. By using a relatively low concentration of reactants, NTP binding becomes the rate-limiting step in elongation.

Kasas *et al* observed RNAPs firmly bound to the surface with associated DNA that was free to diffuse laterally on the surface (Kasas *et al* 1997a). On addition of NTPs, DNA molecules were observed being pulled through the enzyme (see figure 9). The associated DNA arms were observed to concomitantly increase and decrease in size, with the template eventually being released from the polymerase. A sequence of images showing intermediates in transcription elongation was taken, and subsequent measurements of the variation of the relative position of the RNAP on the template with time, allowed the transcription rate to be estimated as being 0.5–2 nt $s^{-1}$. Similar measurements were performed by Guthold *et al* with a fragment containing the $\lambda P_R$ promoter and $t_{R2}$ terminator (Guthold *et al* 1999). Before addition of NTPs, the arms of DNA either side of the polymerase were observed to diffuse laterally but never translocated through the enzyme. Upon nucleotide injection, RNAP began to thread the template in a unidirectional manner, consistent with the

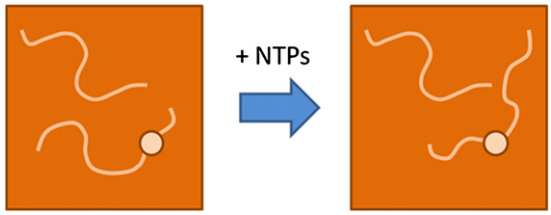


**Figure 9.** Schematic of how dynamic studies into transcription elongation appear under AFM. One DNA strand is depicted with an RNAP bound, whilst the other strand can be used as a reference. After NTPs are added into the reaction mix the RNAP remained stationary, but was observed to thread the DNA through its core in a direction corresponding to transcription.

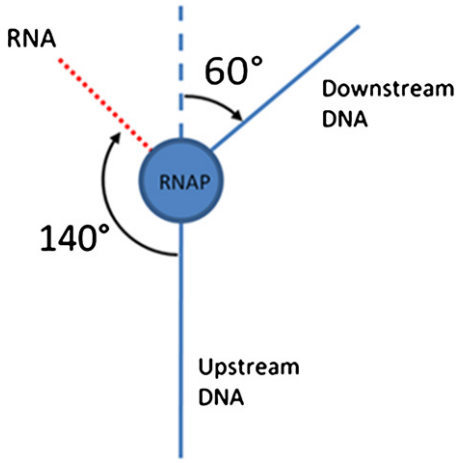


**Figure 10.** Model of transcription elongation complex proposed by Rivetti *et al* (2003).

orientation of the promoter in the fragment. Measurements yielded a transcription rate of 1.5 ± 0.8 nt $s^{-1}$.

Guthold *et al* also performed a biochemical assay in the same buffer conditions as those used in their AFM studies, and found that transcription proceeded at a much faster rate (5 nt $s^{-1}$), compared to the 1.5 nt $s^{-1}$ rate seen with AFM (Guthold *et al* 1999). Again, this can be attributed to the presence of the mica surface in the AFM studies. During transcription, positive supercoils build up in front of RNAP and negative ones behind it (Liu and Wang 1987). In solution these can be readily relaxed by rotation of the strands in space. However, in surface-bound complexes the rotation of the molecules is more constrained and the dissipation of supercoiling is more difficult. This can lead to a build-up of torsional stress in the molecule, which can supply a drag force on RNAP and slow down transcription.

In both these two dynamic studies of transcription it was not possible to observe the RNA transcripts. RNA is able to form base pairs with itself, allowing it to fold up into compact secondary structures. These should remain in contact with RNAP and are difficult to resolve with the AFM tip. However, it has been possible to confirm transcription activity with AFM, by use of rolling circle complexes (Hansma *et al* 1999, Kasas *et al* 1997a). These are small, circular fragments of single-stranded DNA, which mimic the transcription bubble formed during transcription initiation. An RNAP is able to transcribe a rolling circle for tens of minutes, or until NTPs

are no longer available in solution. Kasas *et al* deposited rolling circle complexes on mica and injected NTPs into the fluid cell continuously to initiate transcription (Kasas *et al* 1997a). The cell was then flushed out, after which the sample was dried. Long pieces of RNA, synthesized from rolling circle templates, were readily imaged by AFM in air. This result shows that even though the complex is bound to a surface, transcription is able to proceed, forming a growing RNA chain.

It is much easier to observe the RNA chain when the sample is dried and imaged in air. Rivetti *et al* studied transcription ternary complexes of *E. coli* RNAP and yeast RNAP III, using the method of nucleotide omission to form SECs (Rivetti *et al* 2003). They discovered that it was possible to observe extended RNA transcripts clearly but only when the stall site was over 370 nucleotides away from the promoter site. Structural information was obtained by mapping the orientation of the RNA transcript with respect to the upstream and downstream DNA arms. In general, the RNA chain was located exiting from the RNAP, on the opposite side of the protein relative to the smaller angle sub-tended by the DNA arms. Measurements of the angle between the RNA chain and upstream DNA yielded different values for the bacteria and for the yeast. For the bacterial RNAP, the distribution was centred about 140°, while this value was 110° in the case of the yeast polymerase. The authors believed that the location of the RNA exit site may help in keeping the transcript far away from the DNA arms, allowing the unhindered growth of the RNA chain (Rivetti *et al* 2003) (see figure 10).

### *Termination*

Transcriptional elongation proceeds until the polymerase encounters a specific base sequence, known as a terminator. Here the polymerase detaches from the DNA template, and the completed RNA molecule is released. Limanskaya and Limanskii used AFM to visualize complexes of bacteriophage T7 RNAP with a DNA template during transcription (Limanskaya and Limanskii 2008). The 1414 bp fragment used contained the T7 promoter and terminator asymmetrically located on the template, such that they could be distinguished by their relative positions. An important characteristic of elongation complexes is that they will rapidly dissociate from the DNA chain upon completion of transcription. In this work, Limanskaya and Limanskii were able to control the reaction conditions to image complexes at the terminator site (Limanskaya and Limanskii 2008). The transcription temperature was reduced from 37 °C to 31 °C, resulting in a decrease in the dissociation rate of the elongation complexes. Large complexes could be observed at the terminator site, with their size suggesting that a number of complexes had stopped one after another after completing elongation. Images were also obtained showing three RNAP molecules bound to the template: one at the promoter, another at the terminator site and a third located in between the two. This result shows that as soon as one RNAP begins elongation, another can bind to the promoter and initiate another round of transcription, and could contribute to the high rate of transcription observed *in vivo*.

### *Nested genes and convergent transcription*

The nested gene is an important and interesting phenomenon and refers to a gene that is located within the boundaries of another gene, often within an intron and in the opposite orientation. Nested genes were first discovered in *Drosophila* flies, when it was shown that the gene *Pcp* encoding pupal cuticle protein was found within an intron of adenosine 3 (*ade3*), lying on the opposite DNA strand (Henikoff *et al* 1986). *Ade3* acts as a housekeeping gene and is involved in basic functions needed for the sustenance of the cell, while cuticle protein gene is associated with the development of tissue. Since then, other nested genes have been discovered, including a number in the human genome (Yu *et al* 2005).

There may be a number of explanations for the existence of nested genes. Gibson *et al* have reviewed the biological implications of nested genes (Gibson *et al* 2005). They suggested that natural selection may favour the presence of nested genes which encode proteins of similar function, in order to enhance a biological response. While nested genes that encode proteins with similar function have been reported, work by Yu *et al* suggests that this is the exception rather than the norm (Yu *et al* 2005). The existence of nested genes may raise the question that they may be co-regulated and transcribed simultaneously with the host gene. For example, the nested genes originally discovered in *Drosophila* are both induced during development. It is also possible that oppositely aligned nested genes may be down-regulated. A RNA polymerase moving along a host gene may experience steric hindrance from another RNAP or transcription factors associated with the nested gene.

A direct consequence of nested genes containing two oppositely aligned promoter sites is the process of convergent transcription (Minchiotti and Dinocera 1991). This involves separate polymerases binding to two different promoter sites, and then beginning to transcribe their respective genes. If the two promoter sites are aligned on opposite DNA strands, the transcribing enzymes will move towards each other, until they become so close that they will begin to affect each other, a situation of transcriptional interference (TI). Gibson *et al* suggested several scenarios for the collision event (Gibson *et al* 2005):

(1) Both RNAPs dissociate from the template resulting in a loss of gene expression.
(2) One RNAP is knocked off resulting in the expression of one gene.
(3) Both RNAPs stall and remain bound to the DNA, resulting in a loss of expression and possible DNA ‘roadblock’.

They also suggested that AFM would be a good tool to visualize what happened when converging polymerases were active on the same DNA template, since it can study the process at a single-molecule level. Later Crampton *et al* undertook AFM studies into convergent transcription (Crampton *et al* 2006a, 2006b). One study utilized a template containing two convergently aligned $\lambda_{PR}$ promoters separated by 338 bp and situated asymmetrically on the DNA template, to enable easy identification of individual RNAPs bound at their respective but identical promoters (Crampton *et al* 2006a). Various stages

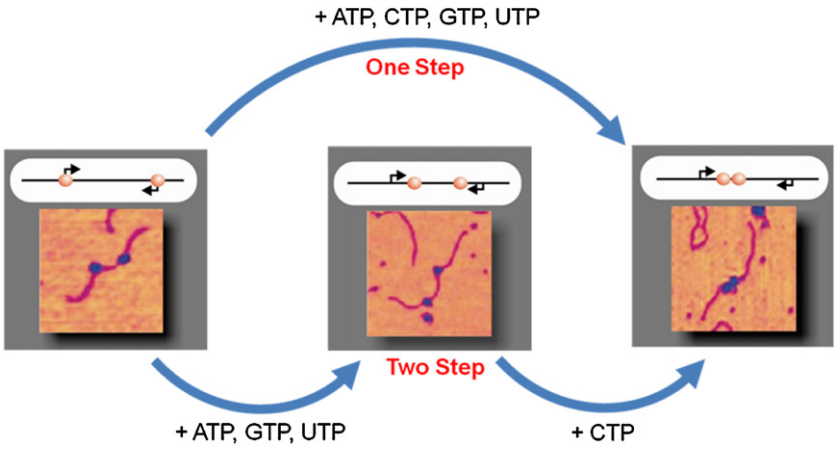


**Figure 11.** Examples of complexes imaged in convergent transcription experiments by Crampton *et al*. The collisions of RNAPs originating from convergently aligned promoters were investigated in a time lapse approach. Collided complexes were formed by two different methods. In the one-step method all four NTPs were added to a mixture containing open promoter complexes. In the two-step approach a stalled intermediate was formed by nucleotide omission. Finally, the missing NTP was added allowing transcription elongation to proceed and caused a collision between the two RNAPs.

of the transcription cycle were imaged (see figure 11 for examples of complexes). Transcription was initiated *in vitro*, and collided complexes were formed in one of two ways. Either stalled intermediates were formed by nucleotide omission, before the remaining NTP was added, or in the other strategy all four NTPs were added in a single step.

A number of outcomes were observed after collision; fragments contained none, one or two polymerase bound. For fragments containing two bound RNAPs, contour length measurements were taken between the apparent centres of each enzyme and also their positions from the closest end of the template. Together these measurements were used in order to infer how the position of each RNAP had changed relative to its starting point.

Of the templates containing two polymerases, a significant number were observed in very close proximity to each other. This was interpreted as the molecules colliding with each other and subsequently stalling. Contour length analysis indicated a second class of complexes. Due to the asymmetric location of the promoter, the DNA exiting each enzyme can either be defined as the short arm or the long arm. In a two-step process whereby transcribing complexes were initially stalled by nucleotide omission and subsequently elongated by adding the missing NTP, contour length distributions taken before and after the injection displayed different distributions over time. The short arm became progressively shorter over the time period studied, whilst the long arm length distribution became broader, with a greater number exhibiting ever longer arm lengths. The interpretation of this data was that one polymerase transcribes into the other and pushes it backwards along the template. Using these data, in conjunction with those of polymerases stalled against each other, it was concluded that the collision induces back-tracking in the other RNAP.

The presence of fragments in these *ex situ* experiments, with just a single or even no polymerases bound does not exclude there being other scenarios. These results could have arisen from the detachment of one or both polymerases after a collision event. This demonstrates the limitations of a time-lapse approach as opposed to real-time imaging. These findings did, however, represent some of the first experimental findings on convergent transcription, and demonstrated the ability of the AFM to give direct structural information, albeit at relatively low resolution (i.e. molecular).

## Conclusions and outlook

AFM has allowed both static and dynamic studies of DNA transcription to be undertaken. The different stages of the transcription process have been studied, both in air and liquid, illustrating the versatility of the technique. These studies have answered some of the questions surrounding transcription and can be used in conjunction with other single-molecule and biochemical techniques to obtain a much clearer picture of the mechanisms involved in transcription. There is plenty of room for development, however, of the AFM technique and the approaches at the single-molecule level. These include: increasing the typically slow scan rates of AFM; increasing understanding of DNA with relevant support surfaces, such as mica; aligning longer DNA molecules to enable study of larger and more complex transcription systems; combining AFM with complementary single-molecule techniques, such as fluorescence microscopy. This final section will discuss advances and future prospects in these different areas.

AFM is especially useful for studying the spatial arrangements of DNA and protein, such as RNAP during different stages of transcription, in order to reveal aspects of biological function. However, a number of technical limitations must be overcome for the technique to be reliably used to visualize dynamic DNA–protein interactions. To obtain any meaningful statistical conclusions about an event, a relatively large population must be studied. Data acquisition is relatively slow with conventional AFM. The imaging time is especially important in dynamic studies, such as transcriptional elongation, where if the tracking is too slow an event may be entirely missed. Scanning a small area does, however, limit the number of molecules visible in a single

scan and as such a large number of scans must be taken of different areas to get a suitable conclusion, a process which can be time-consuming.

Attempts have been made at improving the scan-speed but these require radical re-design of the AFM instrumentation. For example, in the high speed VideoAFM a micro-resonant scanner is used that moves the sample relative to the tip (Humphris *et al* 2003, Hobbs *et al* 2005). As its name suggests, the VideoAFM allows real-time images to be obtained at video frame rates. An image on a conventional AFM takes tens of seconds to obtain, whereas the VideoAFM can deliver videos at a rate of around 20 frames per second. There is no electronic feedback loop, but rather it uses passive mechanical feedback. By using a very soft cantilever and applying a down-force directly to the tip, it is made to track the sample surface. The cantilever is, however, susceptible to hydrodynamic interactions of squeeze film damping between the lever and sample, especially due to the fast relative motion of the probe across the sample. This has a tendency to blur out fine detail, as the tip does not reliably track the sample at shorter length scales. More development of the technique is required to enable this technology to be applied to studying biomolecular samples and DNA transcription. In terms of increased throughput for single-molecule studies, however, it shows promise for imaging DNA samples in ambient air conditions that are tightly bound to a mica surface.

Another strategy for high-speed AFM utilizes small cantilevers (Ando *et al* 2001). Miniaturizing AFM components means that their resonant frequencies increase, allowing them to be driven at higher speeds before mechanical resonances are excited. A number of studies utilizing the small cantilever approach have centred on the dynamics of restriction nucleases and DNA complexes at a single-molecule level, during which it was possible to study the interplay between the two molecules, such as the one-dimensional diffusion of a restriction enzyme along DNA, and the eventual cleavage of the strand in real time (Crampton *et al* 2007, Gilmore *et al* 2009, Yokokawa *et al* 2005). One such study by Crampton *et al* studied the action of the type III restriction nuclease EcoP151 (Crampton *et al* 2007), which must interact with two recognition sites separated by 3500 bp before it is able to cleave the DNA. The fast-scanning method allowed the imaging of EcoP151 translocation in real time at a rate of 79 bp $s^{-1}$, and a scan rate 1–3 frames per second. This is a significant break-through, as illustrating by comparing this imaging rate to the speed of translocation of RNAP on DNA observed by Guthold *et al* using conventional tapping AFM (Guthold *et al* 1999). They also observed the dynamic formation of DNA loops and transient DNA super-coiling. They concluded that the enzyme uses both translocation and diffusive looping to form contacts with its recognition sites.

Small cantilever AFMs using tapping mode have also been used by Kobayashi *et al* to visualize short DNA strands in motion in aqueous liquid (Kobayashi *et al* 2007). Moreover, they demonstrated the ability of the technique to observe real-time images of biotinylated DNA binding to and dissociating from streptavidin protein. It has already been demonstrated that it is possible to form DNA–RNAP complexes in bulk fluid using conventional tapping-mode AFM by injecting a solution of RNAP holoenzyme into the fluid cell (Guthold *et al* 1994). As such, it should be possible to observe the real-time formation of OPCs using fast-scan AFM techniques. Fast-scan AFM still holds the promise of being able to study DNA transcription *in situ*, with a view to observing the process in its entirety from promoter recognition to termination.

The transverse dynamic force microscope (TDFM), also called shear force microscopy, uses a cylindrical probe that is mounted perpendicular to the sample, which is set into transverse oscillation (Antognozzi *et al* 2001, 2003, Humphris *et al* 2002). This technique can be operated using non-contact methods, whereby the probe detects the sample surface at about 1 nm separation, and image contrast is a result of the change in mechanical properties of the water layer confined between the probe tip and the sample (Antognozzi *et al* 2003). Using this technique very low forces are applied to the sample, resulting in negligible sample distortion, particularly important when studying biological samples. When compared with conventional tapping-mode AFM, measured heights of DNA molecules were observed to be higher when imaged with TDFM (Antognozzi *et al* 2002). The technique is currently being developed that would allow images to be captured at ultra-high speed in a liquid environment, allowing biological process to be followed with millisecond time resolution.

It has already been shown that it is possible to study DNA–protein dynamics with fast-scan AFMs (Crampton *et al* 2007). However the technique is in its infancy, and more work is needed to make the system reliable. Even at these enhanced scan rates, resolution is compromised due to molecular motion and routine imaging of DNA is difficult to achieve because tuning of the DNA–mica interaction is often inconsistent. However, work on understanding how divalent cations bind DNA to a mica surface through counter-ion correlations has allowed a greater degree of control over DNA conformations in buffer (Pastre *et al* 2003, 2006). Further advances in the understanding of DNA binding to mica will help to optimize sample preparation for fast-scan AFM experiments.

To date, studies of DNA transcription with AFM have generally used fairly simple model systems, typically consisting of short DNA templates containing the required promoter regions, and RNAP holoenzymes. Transcription *in vivo* is much more complex; genomic DNA can be thousands of base pairs long, and a RNAP may require a number of different transcription factors to assist activation of transcription. For further insights into transcription to be gained through AFM techniques, it will become necessary to switch to systems of greater complexity, which better imitate the conditions found *in vivo*.

A major problem when imaging long strands of DNA with AFM is their tendency to become tangled on the surface during deposition. It would be difficult to study transcription along a large piece of genomic DNA using AFM, as typical sample preparation methods would lead to a highly tangled molecule, meaning that regions of the chain would be obscured and accurate measurements would be difficult to obtain. One method for aligning long DNA molecules on a surface uses the flow force generated by meniscus motion, otherwise known as

molecular combing. Li *et al* used gas flow to drive forward a drop of DNA solution on a bare mica surface (Li *et al* 1998). By carefully controlling the direction of the flow, and the speed of the moving interface they were able to align long λ–DNA fragments (48 502 bp) on the surface. Maaloum pursued a different strategy using lipid monolayers as a substrate for DNA (Maaloum 2003). Local de-mixing of the different lipid constituents caused λ–DNA to bind to the surface in a preferred orientation. The favourable orientation of the DNA relative to the fast-scan direction allowed high resolution images to be collected under aqueous buffer, showing an axial repeat consistent with the helical pitch of DNA.

The work of Crampton *et al* on convergent transcription represents one the first studies where more than one protein has been observed on a single DNA template (Crampton *et al* 2006a, 2006b). To study systems of greater and greater complexity to further understand transcription *in vivo*, new methods complementary to AFM will be necessary, for example, to identify different protein components within a larger macromolecular complex. Combining AFM with fluorescence microscopy is an attractive solution, since specific molecules can be identified by attaching fluorescent dyes or objects. However, the spatial resolution of the technique is limited by the wavelength of light used during excitation. By combining both techniques together it becomes possible to overcome their own individual limitations. Ebenstein *et al* used this strategy to examine the interaction of two distinct RNAPs (*E. coli* RNAP and T7 RNAP) with T7 genomic DNA (Ebenstein *et al* 2009). The DNA possessed promoter sites for both RNAPs. Three *E. coli* RNAPs were observed with AFM in close proximity to each other. The *E. coli* RNAPs were labelled with a fluorescent quantum dot, and whilst individual molecules could not be resolved, the fluorescence image identified the particular type of protein observed under AFM. The T7 RNAPs were labelled with four different coloured quantum dots, which could be seen spaced along the DNA backbone with fluorescence imaging, demonstrating the ability to detect multiple interacting species.

Sanchez *et al* used fluorescent polystyrene nano-spheres to align the optical image with the AFM topographic image, allowing both signals to be overlaid for easy analysis (Sanchez *et al* 2010). They studied the interaction of the human RAD51 protein labelled with a single fluorophore with λ–DNA. Under fluorescence imaging they observed green filamentous structures, which were interpreted as DNA covered with RAD51. In this case they were able to detect the presence of the protein by its fluorescence, and further define the DNA structure from the AFM topography images. Although single-fluorophore sensitivity has not yet been demonstrated by this technique, it is nonetheless an intriguing development in imaging, combining the single-molecular resolution of AFM with molecular recognition capabilities from optical signals.

Over the last couple of decades, AFM has been used to study every major step in transcription, from the initial formation of DNA–RNAP complexes to dynamic studies of transcription elongation, and eventual termination of complexes. The microscopes ability to provide direct structural information has proved extremely useful in ascertaining the spatial arrangements of DNA–RNAP complexes actively involved in transcription. The potential future developments in AFM imaging described above demonstrate that the technique will continue to reveal new insights into the basic mechanisms of transcription, and will allow the possibility of moving away from predominantly simple studies on static systems to dynamic systems containing a number of interacting proteins, more akin to the situations found *in vivo*.

## References


Alessandrini A and Facci P 2005 AFM: a versatile tool in biophysics *Meas. Sci. Technol.* **16** R65–92

Ando T, Kodera N, Takai E, Maruyama D, Saito K and Toda A 2001 A high-speed atomic force microscope for studying biological macromolecules *Proc. Natl Acad. Sci. USA* **98** 12468–72

Ansari A Z 2007 Chemical crosshairs on the central dogma *Nat. Chem. Biol.* **3** 2–7

Antognozzi M, Humphris A D L and Miles M J 2001 Observation of molecular layering in a confined water film and study of the layers viscoelastic properties *Appl. Phys. Lett.* **78** 300–2

Antognozzi M, Szczelkun M D, Humphris A D L and Miles M J 2003 Increasing shear force microscopy scanning rate using active quality-factor control *Appl. Phys. Lett.* **82** 2761–3

Antognozzi M, Szczelkun M D, Round A N and Miles M J 2002 Comparison between shear force and tapping mode AFM—high resolution imaging of DNA *Single Mol.* **3** 105–10

Bai L, Santangelo T J and Wang M D 2006 Single-molecule analysis of RNA polymerase transcription *Annu. Rev. Biophys. Biomol. Struct.* **35** 343–60

Bates A D and Maxwell A 1993 *DNA Topology* (Oxford: IRL Press)

Berg O G, Winter R B and von Hippel P H 1981 Diffusion-driven mechanisms of protein translocation on nucleic-acids: 1. Models and theory *Biochemistry (Mosc.)* **20** 6929–48

Billingsley D J, Kirkham J, Bonass W A and Thomson N H 2010 Atomic force microscopy of DNA at high humidity: irreversible conformational switching of supercoiled molecules *Phys. Chem. Chem. Phys.* **12** 14727–34

Bustamante C, Guthold M, Zhu X S and Yang G L 1999 Facilitated target location on DNA by individual Escherichia coli RNA polymerase molecules observed with the scanning force microscope operating in liquid *J. Biol. Chem.* **274** 16665–8

Bustamante C, Vesenka J, Tang C L, Rees W, Guthold M and Keller R 1992 Circular DNA-molecules imaged in air by scanning force microscopy *Biochemistry (Mosc.)* **31** 22–6

Carpousis A J and Gralla J D 1980 Cycling of ribonucleic-acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter *Biochemistry (Mosc.)* **19** 3245–53

Corden J, Wasylyk B, Buchwalder A, Corsi P S, Kedinger C and Chambon P 1980 Promoter sequences of eukaryotic protein-coding genes *Science* **209** 1406–14

Crampton N, Bonass W A, Kirkham J, Rivetti C and Thomson N H 2006a Collision events between RNA polymerases in convergent transcription studied by atomic force microscopy *Nucleic Acids Res.* **34** 5416–25

Crampton N, Bonass W A, Kirkham J and Thomson N H 2005 Formation of aminosilane-functionalized mica for atomic force microscopy imaging of DNA *Langmuir* **21** 7884–91

Crampton N, Thomson N H, Kirkham J, Gibson C W and Bonass W A 2006b Imaging RNA polymerase-amelogenin gene complexes with single molecule resolution using atomic force microscopy *Eur. J. Oral Sci.* **114** 133–8

Crampton N, Yokokawa M, Dryden D T F, Edwardson J M, Rao D N, Takeyasu K, Yoshimura S H and Henderson R M 2007 Fast-scan atomic force microscopy reveals that the type III restriction enzyme EcoP151 is capable of DNA

translocation and looping *Proc. Natl Acad. Sci. USA* **104** 12755–60
Dame R T, Wyman C and Goosen N 2003 Insights into the regulation of transcription by scanning force microscopy *J. Microsc. (Oxford)* **212** 244–53
Darst S A, Kubalek E W and Kornberg R D 1989 Three-dimensional structure of Escherichia coli RNA polymerase holoenzyme determined by electron crystallography *Nature* **340** 730–2
Das A 1993 Control of transcription termination by RNA-binding proteins *Annu. Rev. Biochem.* **62** 893–930
deHaseth P L, Zupancic M L and Record M T 1998 RNA polymerase-promoter interactions: the comings and goings of RNA polymerase *J. Bacteriol.* **180** 3019–25
Durniak K J, Bailey S and Steitz T A 2008 The structure of a transcribing T7 RNA polymerase in transition from initiation to elongation *Science* **322** 553–7
Ebenstein Y, Gassman N, Kim S and Weiss S 2009 Combining atomic force and fluorescence microscopy for analysis of quantum-dot labeled protein–DNA complexes *J. Mol. Recognit.* **22** 397–402
Garcia R and Perez R 2002 Dynamic atomic force microscopy methods *Surf. Sci. Rep.* **47** 197–301
Gibson C W, Thomson N H, Abrams W R and Kirkham J 2005 Nested genes: biological implications and use of AFM for analysis *Gene* **350** 15–23
Gilmore J L, Suzuki Y, Tamulaitis G, Siksnys V, Takeyasu K and Lyubchenko Y L 2009 Single-molecule dynamics of the DNA-EcoRII protein complexes revealed with high-speed atomic force microscopy *Biochemistry (Mosc.)* **48** 10492–8
Guthold M, Bezanilla M, Erie D A, Jenkins B, Hansma H G and Bustamante C 1994 Following the assembly of RNA polymerase-DNA complexes in aqueous solutions with the scanning force microscope *Proc. Natl Acad. Sci. USA* **91** 12927–31
Guthold M, Zhu X, Rivetti C, Yang G, Thomson N H, Kasas S, Hansma H G, Smith B, Hansma P K and Bustamante C 1999 Direct observation of one-dimensional diffusion and transcription by Escherichia coli RNA polymerase *Biophys. J.* **77** 2284–94
Hansma H G 2001 Surface biology of DNA by atomic force microscopy *Annu. Rev. Phys. Chem.* **52** 71–92
Hansma H G, Golan R, Hsieh W, Daubendiek S L and Kool E T 1999 Polymerase activities and RNA structures in the atomic force microscope *J. Struct. Biol.* **127** 240–7
Hansma H G, Weisenhorn A L, Gould S A C, Sinsheimer R L, Gaub H E, Stucky G D, Zaremba C M and Hansma P K 1991 Progress in sequencing deoxyribonucleic-acid with an atomic force microscope *J. Vac. Sci. Technol.* B **9** 1282–4
Hansma P K, Cleveland J P, Radmacher M, Walters D A, Hillner P E, Bezanilla M, Fritz M, Vie D and Hansma H G 1994 Tapping mode atomic force microscopy in liquids *Appl. Phys. Lett.* **64** 1738–40
Henikoff S, Keene M A, Fechtel K and Fristrom J W 1986 Gene within a gene–nested drosophila genes encode unrelated proteins on opposite DNA strands *Cell* **44** 33–42
Herbert K M, Greenleaf W J and Block S M 2008 Single-molecule studies of RNA polymerase: motoring along *Annu. Rev. Biochem.* **77** 149–76
Hobbs J K, Vasilev C and Humphris A D L 2005 Real time observation of crystallization in polyethylene oxide with video rate atomic force microscopy *Polymer* **46** 10226–36
Hsu L M 2002 Promoter clearance and escape in prokaryotes *Biochim. Biophys. Acta* **1577** 191–207
Hu J, Xiao X D, Ogletree D F and Salmeron M 1995 The structure of molecularly thin films of water on mica in humid environments *Surf. Sci.* **344** 221–36
Humphris A D L, Antognozzi M, McMaster T J and Miles M J 2002 Transverse dynamic force spectroscopy: a novel approach to determining the complex stiffness of a single molecule *Langmuir* **18** 1729–33
Humphris A D L, Hobbs J K and Miles M J 2003 Ultra high-speed scanning probe microscopy capable of over 100 frames per second *Scanning Tunneling Microscopy/Spectroscopy and Related Techniques AIP Conf. Proc.* **696** 79–85
Hurwitz J 2005 The discovery of RNA polymerase *J. Biol. Chem.* **280** 42477–85
Hurwitz J, Bresler A and Diringer R 1960 The enzymic incorporation of ribonucleotides into polyribonucleotides and the effect of DNA *Biochem. Biophys. Res. Commun.* **3** 15–9
Kasas S 2003 Atomic Force Microscopy: Biomedical methods and applications *Sample preparation method for observing RNA polymerase activity by atomic force microscopy* (Totowa, NJ: Humana Press) chapter 25
Kasas S, Thomson N H, Smith B L, Hansma H G, Zhu X S, Guthold M, Bustamante C, Kool E T, Kashlev M and Hansma P K 1997a Escherichia coli RNA polymerase activity observed using atomic force microscopy *Biochemistry (Mosc.)* **36** 461–8
Kasas S, Thomson N H, Smith B L, Hansma P K, Miklossy J and Hansma H G 1997b Biological applications of the AFM: from single molecules to organs *Int. J. Imaging Syst. Technol.* **8** 151–61
Kobayashi M, Sumitomo K and Torimitsu K 2007 Real-time imaging of DNA-streptavidin complex formation in solution using a high-speed atomic force microscope *Ultramicroscopy* **107** 184–90
Lang D and Coates P 1968 Diffusion coefficient of DNA in solution at 'Zero' concentration as measured by electron microscopy *J. Mol. Biol.* **36** 137–51
Li J W, Bai C L, Wang C, Zhu C F, Lin Z, Li Q and Cao E H 1998 A convenient method of aligning large DNA molecules on bare mica surfaces for atomic force microscopy *Nucleic Acids Res.* **26** 4785–6
Limanskaya O Y and Limanskii A P 2008 Imaging of T7 RNA polymerase elongation complexes by atomic force microscopy *Mol. Biol.* **42** 469–77
Lindsay S M, Nagahara L A, Thundat T, Knipping U, Rill R L, Drake B, Prater C B, Weisenhorn A L, Gould S A C and Hansma P K 1989 STM and AFM images of nucleosome DNA under water *J. Biomol. Struct. Dyn.* **7** 279–87
Liu L F and Wang J C 1987 Supercoiling of the DNA-template during transcription *Proc. Natl Acad. Sci. USA* **84** 7024–7
Lyubchenko Y, Shlyakhtenko L, Harrington R, Oden P and Lindsay S 1993 Atomic force microscopy of long DNA—imaging in air and under water *Proc. Natl Acad. Sci. USA* **90** 2137–40
Maaloum M 2003 A close encounter with DNA *Eur. Biophys. J.* **32** 585–7
McClure W R 1985 Mechanism and control of transcription initiation in prokaryotes *Annu. Rev. Biochem.* **54** 171–204
Meyer G and Amer N M 1988 Novel optical approach to atomic force microscopy *Appl. Phys. Lett.* **53** 1045–7
Minchiotti G and Dinocera P P 1991 Convergent transcription initiates from oppositely oriented promoters within the 5′ end regions of Drosophila melanogaster F elements *Mol. Cell. Biol.* **11** 5171–80
Pastre D, Hamon L, Landousy F, Sorel I, David M O, Zozime A, Le Cam E and Pietrement O 2006 Anionic polyelectrolyte adsorption on mica mediated by multivalent cations: a solution to DNA imaging by atomic force microscopy under high ionic strengths *Langmuir* **22** 6651–60
Pastre D, Pietrement O, Fusil P, Landousy F, Jeusset J, David M O, Hamon C, Le Cam E and Zozime A 2003 Adsorption of DNA to mica mediated by divalent counterions: a theoretical and experimental study *Biophys. J.* **85** 2507–18

Polyakov A, Severinova E and Darst S A 1995 Three-dimensional structure of E. coil core RNA polymerase: promoter binding and elongation conformations of the enzyme *Cell* **83** 365–73

Putman C A J, Van der Werf K O, De Grooth B G, Van Hulst N F and Greve J 1994 Tapping mode atomic force microscopy in liquid *Appl. Phys. Lett.* **64** 2454–6

Rees W A, Keller R W, Vesenka J P, Yang G L and Bustamante C 1993 Evidence of DNA bending in transcription complexes imaged by scanning force microscopy *Science* **260** 1646–9

Ring B Z, Yarnell W S and Roberts J W 1996 Function of E. coli RNA polymerase [sigma] factor- [sigma]70 in promoter-proximal pausing *Cell* **86** 485–93

Rippe K, Guthold M, von Hippel P H and Bustamante C 1997 Transcriptional activation via DNA-looping: visualization of intermediates in the activation pathway of E-coli RNA polymerase center dot sigma(54) holoenzyme by scanning force microscopy *J. Mol. Biol.* **270** 125–38

Rivetti C, Codeluppi S, Dieci G and Bustamante C 2003 Visualizing RNA extrusion and DNA wrapping in transcription elongation complexes of bacterial and eukaryotic RNA polymerases *J. Mol. Biol.* **326** 1413–26

Rivetti C, Guthold M and Bustamante C 1996 Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis *J. Mol. Biol.* **264** 919–32

Rivetti C, Guthold M and Bustamante C 1999 Wrapping of DNA around the E. coli RNA polymerase open promoter complex *EMBO J.* **18** 4464–75

Sanchez H, Kanaar R and Wyman C 2010 Molecular recognition of DNA-protein complexes: a straightforward method combining scanning force and fluorescence microscopy *Ultramicroscopy* **110** 844–51

Schulz A, Mucke N, Langowski J and Rippe K 1998 Scanning force microscopy of Escherichia coli RNA polymerase. Sigma(54) holoenzyme complexes with DNA in buffer and in air *J. Mol. Biol.* **283** 821–36

Shlyakhtenko L S, Gall A A, Weimer J J, Hawn D D and Lyubchenko Y L 1999 Atomic force microscopy imaging of DNA covalently immobilized on a functionalized mica substrate *Biophys. J.* **77** 568–76

Shlyakhtenko L S, Potaman V N, Sinden R R and Lyubchenko Y L 1998 Structure and dynamics of supercoil-stabilized DNA cruciforms *J. Mol. Biol.* **280** 61–72

Singer P and Wu C W 1987 Promoter search by Escherichia-coli RNA-polymerase on a circular DNA-template *J. Biol. Chem.* **262** 14178–89

Stevens A 1960 Incorporation of the adenine ribonucleotide into RNA by cell fractions from, E-coli B *Biochem. Biophys. Res. Commun.* **3** 92–6

Thomson N H 2006 *Applied Scanning Probe Methods. Characterization* ed B Bhushan *et al* (Berlin: Springer) vol VI

Thomson N H, Kasas S, Smith B, Hansma H G and Hansma P K 1996 Reversible binding of DNA to mica for AFM imaging *Langmuir* **12** 5905–8

Thundat T, Warmack R J, Allison D P, Bottomley L A, Lourenco A J and Ferrell T L 1992 Atomic force microscopy of deoxyribonucleic acid strands adsorbed on mica: the effect of humidity on apparent width and image contrast *J. Vac. Sci. Technol.* A **10** 630–5

Vesenka J, Guthold M, Tang C L, Keller D, Delaine E and Bustamante C 1992 Substrate preparation for reliable imaging of DNA molecules with the scanning force microscope *Ultramicroscopy* **42–44** 1243–9

von Hippel P H and Berg O G 1989 Facilitated target location in biological systems *J. Biol. Chem.* **264** 675–8

Wiggins P A, van der Heijden T, Moreno-Herrero F, Spakowitz A, Phillips R, Widom J, Dekker C and Nelson P C 2006 High flexibility of DNA on short length scales probed by atomic force microscopy *Nat. Nano* **1** 137–41

Yokokawa M, Takeyasu K and Yoshimura S H 2005 Fast-scanning AFM is now applicable to the analyses of single-molecule reactions in nano-biophysics *2005 IEEE Int. Symp. on Micro-NanoMechatronics and Human Science* pp 121–6

Yu P, Ma D L and Xu M X 2005 Nested genes in the human genome *Genomics* **86** 414–22

Zhang G Y, Campbell E A, Minakhin L, Richter C, Severinov K and Darst S A 1999 Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 angstrom resolution *Cell* **98** 811–24

Zhong Q, Inniss D, Kjoller K and Elings V B 1993 Fractured polymer/silica fiber surface studied by tapping mode atomic force microscopy *Surf. Sci.* **290** L688–92