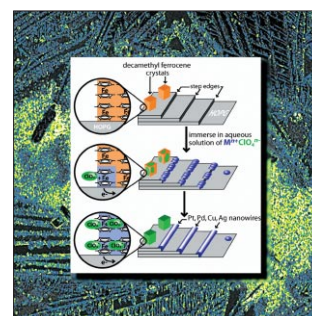
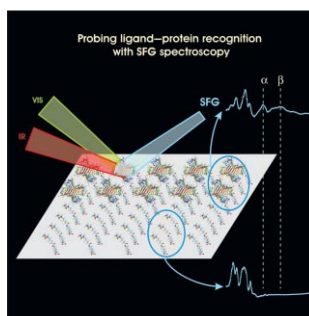
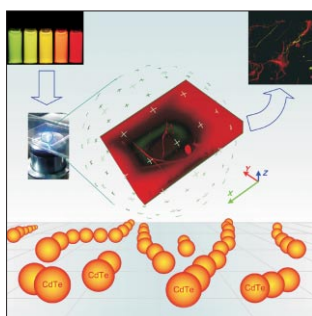
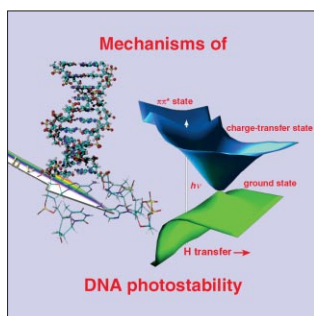
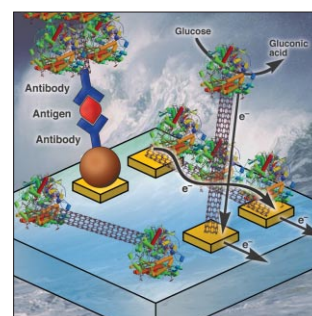
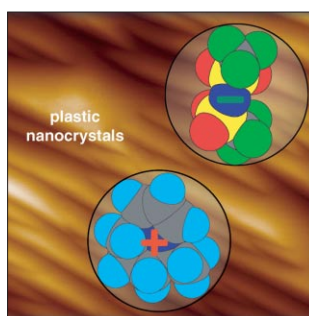
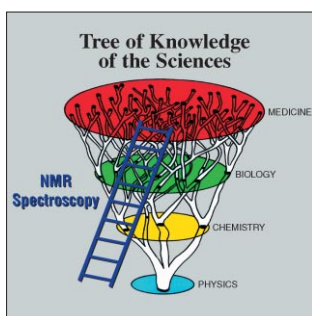
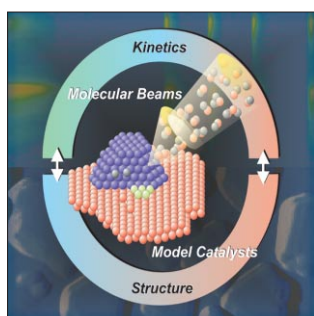
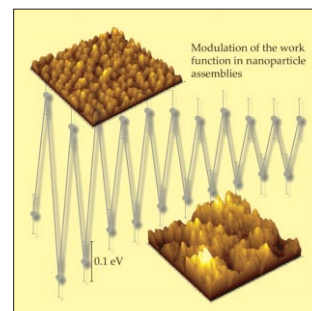
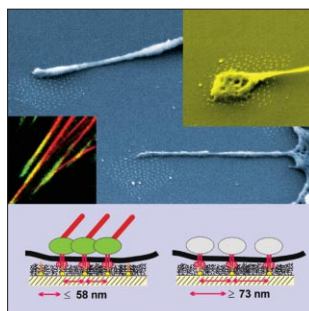
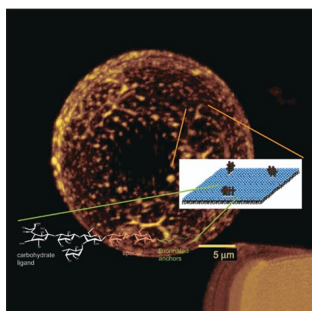
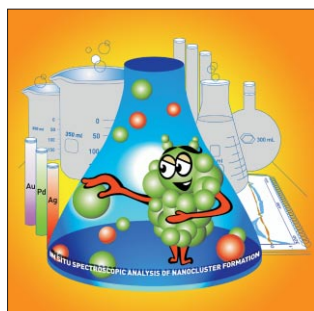


A EUROPEAN JOURNAL

# CHEMPHYSICHEM

OF CHEMICAL PHYSICS AND PHYSICAL CHEMISTRY



## Reprint

© Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

 WILEY-VCH

A EUROPEAN JOURNAL

# CHEMPHYSICHEM

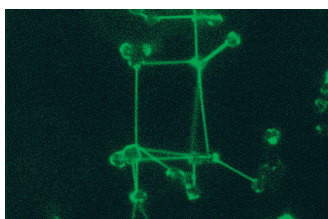
OF CHEMICAL PHYSICS AND PHYSICAL CHEMISTRY

## Table of Contents

*K. O. Greulich\**

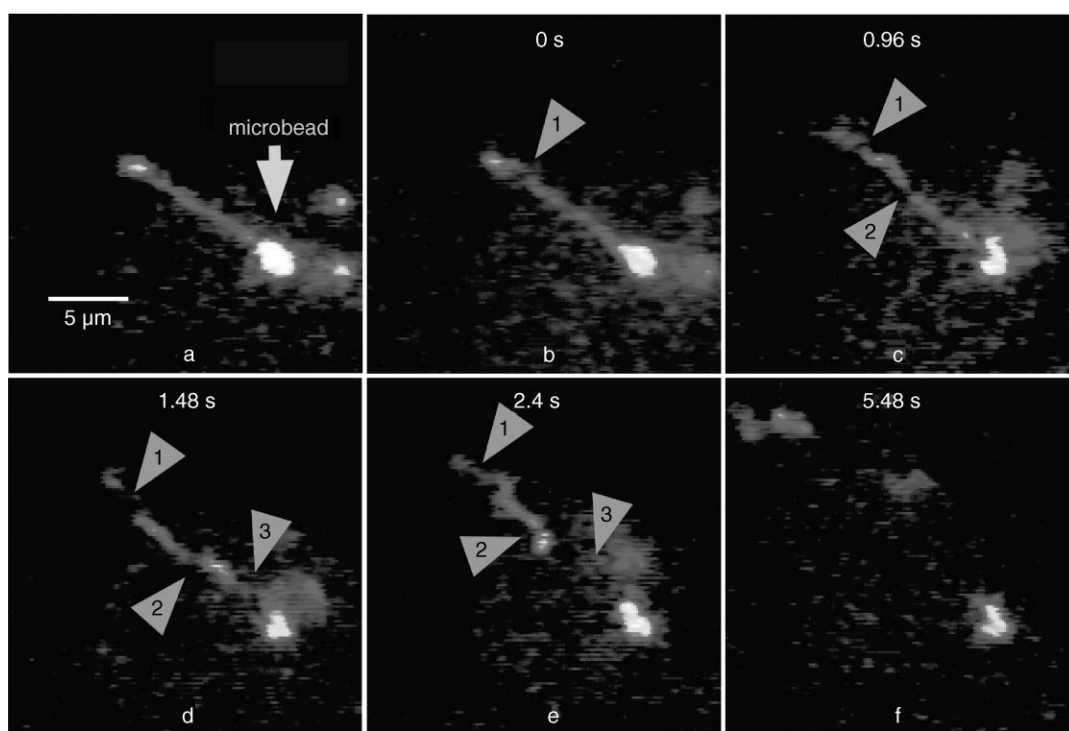
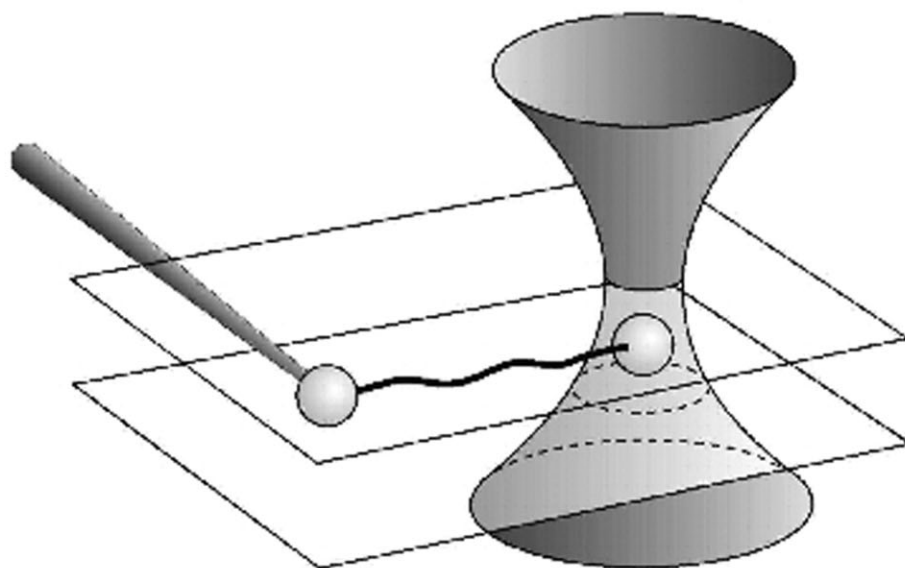
2458 – 2471

**Single-Molecule Studies on DNA and RNA**

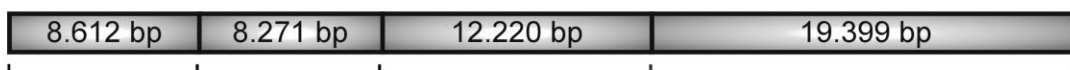


**Single file:** DNA and RNA are molecules with a high degree of individuality, and therefore single-molecule techniques are necessary to understand their properties in sufficient detail (see picture). This Review examines recent techniques for the preparation, handling, and detection of single molecules of nucleic acids, as well as some of the obtained results. These include the construction of nanomachines, such as a walking single-molecule robot and a single-molecule DNA computer.

## Fluorescence spectroscopy on single biomolecules



$\lambda$ -DNA (48.502 bp) + Sma I:



# Single-Molecule Studies on DNA and RNA

Karl Otto Greulich\*<sup>[a]</sup>

*DNA and RNA are the most individual molecules known. Therefore, single-molecule experiments with these nucleic acids are particularly useful. This review reports on recent experiments with single DNA and RNA molecules. First, techniques for their preparation and handling are summarised including the use of AFM nanotips and optical or magnetic tweezers. As important detection techniques, conventional and near-field microscopy as well as fluorescence resonance energy transfer (FRET) and fluorescence correlation spectroscopy (FCS) are touched on briefly. The use of single-molecule techniques currently includes force measurements in stretched nucleic acids and in their complexes with binding partners, particularly proteins, and the analysis of DNA*

*by restriction mapping, fragment sizing and single-molecule hybridisation. Also, the reactions of RNA polymerases and enzymes involved in DNA replication and repair are dealt with in some detail, followed by a discussion of the transport of individual nucleic acid molecules during the readout and use of genetic information and during the infection of cells by viruses. The final sections show how the enormous addressability in nucleic acid molecules can be exploited to construct a single-molecule field-effect transistor and a walking single-molecule robot, and how individual DNA molecules can be used to assemble a single-molecule DNA computer.*

## 1. Introduction

Single-biomolecule research started as early as 1961, when a single  $\beta$ -galactosidase molecule in a microdroplet was studied.<sup>[1]</sup> In 1995, a first special issue of a regular journal on "single-molecule detection" appeared.<sup>[2]</sup> A few recent reviews have at least attempted to cover the field of single-biomolecule research as a whole.<sup>[3–5]</sup> A review dealing primarily with single DNA molecule experiments is available from Ishi and Yanagida.<sup>[6]</sup> Meanwhile, single-biomolecule research is moving at such a fast pace that it is impossible to cover this topic in one review article. Therefore, the present overview is restricted to single DNA and RNA molecule research. Even with this restriction, there is only room for a report on selected, recently published work. Therefore, this review is essentially a continuation of other reviews.<sup>[3–5]</sup> The work discussed there, and older work, has only been carried over into the present review when it either gives the reader a good introduction to the respective subject, or when it complements and completes work discussed here. The vast majority of the papers discussed appeared in the latter half of 2002 or more recently.

DNA and RNA are probably the molecules with the highest degree of individuality, even more than that of proteins, which, in this respect, represent a subclass. How far this individuality goes is illustrated in the following two examples (for more details, see reference [7]).

The double-stranded DNA of each human cell contains  $3 \times 10^9$  base pairs, and the distance from base pair to base pair is 3 Å (0.3 nm). Multiplication of both figures gives a length of approximately 1 m. This has to be doubled, since each human cell contains at least two sets of DNA, one inherited from the mother, the other from the father. Thus, as an estimate, we have:

- 1) two metres of DNA per human cell. The human body consists of more than  $10^{13}$  cells. Multiplication of the number of cells by the DNA length gives us:
- 2)  $2 \times 10^{13}$  m DNA in the body of each human individual. Given the fact that 1 080 000 000 km ( $\approx 1$  light hour) is the distance between the Earth and the Saturn moon, Titan, which a spacecraft needed 7 years to cover, this corresponds to 18.5 such distances or 0.77 light days. In other words we are faced with the following surprising equivalence:
- 3) the DNA of approximately 500 human individuals spans one light year. While the length of DNA molecules is in fact astronomical, their diameter is small: approximately 2 nm. If the sequence variability of DNA is required there is a second surprise: the principal variability of DNA, given the length of human DNA, is far beyond anything one can imagine. Therefore, a much simpler question should be asked: how many different DNA molecules with a given length of only 120 base pairs are possible, and how much material do we need to synthesise just one "exemplar" of all of these molecules? For molecules of length 1 there are four possibilities: A C T G. For molecules of length 2 there are  $4^2 = 16$  possibilities: AA, AC, AG, AT, CA, CC...
- 4) For all molecules with a length of 120 base pairs there are  $4^{120} = 10^{72}$  possibilities. Since 1 mol is  $6 \times 10^{23}$  molecules, this corresponds to  $1.6 \times 10^{48}$  mol. The molecular mass of a base pair is 700 g, therefore molecules with 120 base pairs

[a] Prof. Dr. K. O. Greulich  
Institute of Molecular Biotechnology  
Beutenbergstr. 11, 07745 Jena (Germany)  
Fax: (+49) 3641-656410  
E-mail: kog@imb-jena.de

have a mass of  $84 \text{ kg mol}^{-1}$ . Thus,  $1.6 \times 10^{48}$  mol corresponds to

- 5)  $100 \times 10^{46} \text{ kg} = 2.5 \times 10^{19}$  masses of the Sun.
- 6) In conclusion, more than the visible mass of the Universe is needed to synthesise a single molecule of each possible variant of a DNA molecule with 120 bases.

These two examples show that DNA and RNA are molecules with a much higher individuality than any other sort of molecule, and therefore single-molecule techniques are more necessary than for any other molecule if we are to understand their properties in sufficient detail. In the following overview it will turn out that many of these techniques come from chemical physics and physical chemistry. Among them are conventional spectroscopic techniques whose sensitivity is increased to the single-molecule limit. Others are derived from sophisticated micromanipulation techniques.

## 2. Materials and Methods: Techniques for Detection and Preparation

A comprehensive overview of the techniques for the preparation and detection of single DNA and RNA molecules would require a review on its own. Therefore, only a few selected techniques are described herein, those which, in the experience of the author, are among the most widely used techniques or are particularly useful for other reasons. The described techniques include the preparation of single DNA molecules and, as a second class of techniques, single-molecule detection approaches.

Karl Otto Greulich studied physics at the University of Heidelberg. From 1980–1982 he worked at the European Molecular Biology Laboratory (EMBL) in Heidelberg on laser fluorescence techniques and from 1982–1983 at the Weizmann Institute of Science in Rehovot, Israel. Back to Heidelberg in 1984 he initiated, together with J. Wolfrum, the working group on "Laser Applications in Biology", where among others optical tweezers were first applied outside the US and complete micromanipulation by light was first developed worldwide. Since 1992 he is department head at the Fritz Lipmann Institute (until August 2005 Institute for Molecular Biotechnology) in Jena and since 1993 also Professor of Biophysics at the Friedrich Schiller University Jena. In 1999, he published the first single author book on "Micromanipulation by light in biology and medicine: The laser microbeam and optical tweezers". From 2000–2002 he was Editor of the journal *Single Molecules* which in 2003 was merged with *ChemPhysChem*, where he is now member of the Editorial Advisory Board.



### 2.1. Preparation and Handling of Single Nucleic Acid Molecules

The most perfect single-molecule DNA or RNA experiment would use an individual molecule prepared from a single cell from a known tissue of a known donor. Only then would it be possible to correlate the molecular individuality with a given phenotype, for example, with a certain disease in a human patient. Such a preparation technique, while it may be envisioned in the future, does not yet exist. Therefore, all presently known single-molecule DNA and RNA preparation techniques start from a bulk of molecules and use a sort of limiting dilution. Only the last step of handling the molecules is a true single-molecule technique. The information on the original donor of the molecule is lost. While this is, in terms of an ideal single-molecule experiment, a significant drawback, it still allows a large number of interesting and informative experiments that are not possible with true bulk techniques.

As already indicated, limiting dilution is one approach to preparing single molecules of any kind, not only DNA and RNA. There is a quite simple relationship between the concentration and the average distance between the molecules [Eq. (1)]:

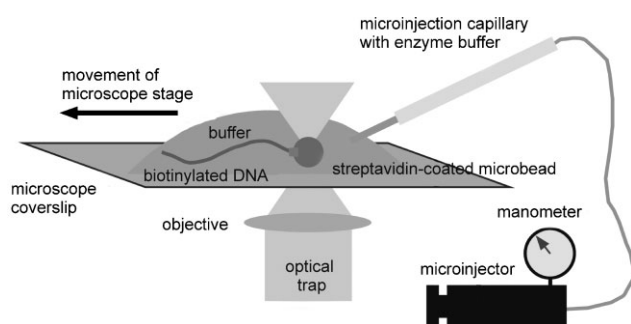
$$d = 1.185/c^{1/3} \quad (1)$$

where  $d$  is the average intermolecular distance in nanometres when  $c$  is given in moles per litre. For example, the intermolecular distance at a concentration of  $1 \text{ mol l}^{-1}$  is 1.185 nm. Correspondingly, for a concentration of  $1 \mu\text{mol l}^{-1}$  this distance is 118.5 nm, and for  $1 \text{ pmol l}^{-1}$  it is 11.8  $\mu\text{m}$ . For a derivation of Equation (1), see references [4] or [8].

A second approach is molecular combing. Here, DNA molecules are embedded in a microdroplet. During evaporation, the meniscus of the droplet on a glass slide recedes. The DNA molecules try to avoid the environment of high surface tension at the borderline between droplet and air, and will in the end be attached to the surface of the glass slide as a highly parallel array of stretched DNA molecules.<sup>[9]</sup>

True handling of individual DNA molecules becomes possible with the tip of an atomic force microscope (AFM), which is essentially a glass capillary pulled to a diameter of 10–20 nm or a very sharp edge of a tetrahedral crystal. The individual DNA molecule is either attached directly by nonspecific adhesion, or the tip itself carries a short piece of a sequence that is complementary to the DNA molecule of interest, and thus hybridises to the latter and binds it. Once bound, the individual DNA molecule can be moved along complex trajectories and stretched with controlled force. The tip of an AFM is best suited to applying and measuring forces ranging from a few hundred piconewtons up to a nanonewton (for comparison, a million of such AFM tips would be required to lift 100 mg on the surface of the earth). The AFM-tip handling technique has opened up a whole field of research, single DNA molecule force spectroscopy. Selected experiments will be reported below.

Finally, individual DNA molecules can be manipulated by using optical tweezers. This approach allows for very gentle handling and sensitive force measurement. Optical tweezers (see, for example, references [7, 10, 11]) are made of a continuous (often near-infrared) laser of moderate power (typically a few hundred milliwatts), which is focused into a microscope to the theoretical (diffraction) limit. Light pressure and gradient forces push and pull microscopic objects towards the focus of the laser. The force exerted depends on the laser power and can be determined by several calibration approaches (see, for example, reference [7]). Since the force transmission between the tweezers and DNA molecule is low, in most cases the DNA molecule is attached to a micrometre-sized polystyrene microbead by techniques similar to those mentioned above for attaching molecules to an AFM tip. Figure 1 shows such a strategy in detail.



**Figure 1.** Attachment to a microbead and manipulation of a single DNA molecule. DNA molecules are end-labelled with biotin using a commercial biotinylation kit. These molecules are fluorescence-stained with SYBRGreen by mixing and stirring. Then the fluorescence-stained, biotinylated molecules are mixed with commercially available streptavidin-coated polystyrene microbeads. By mixing at a suitable stoichiometry, single DNA molecules, or occasionally even two or three, are attached to the beads. The latter are held by optical tweezers, while the environment is moved by moving the xy stage of the microscope, thus stretching the DNA molecule under observation. Enzyme solution is added with a micropipette to start a reaction.

As with AFM tips, the mechanical properties of biomolecules can be measured with optical tweezers. For some applications, magnetic tweezers can be used instead of optical tweezers.<sup>[12]</sup> In this case, the microbeads are magnetic and the force is exerted by magnetic fields. Magnetic tweezers are somewhat less versatile with respect to the trajectories on which a single DNA molecule can be moved, but the range of forces is much larger and application of a torque to a single molecule is possible in a more convenient way than with optical tweezers. Thus, magnetic tweezers are often used to pull the two complementary strands of DNA apart or to stretch and overstretch single DNA molecules. Selected experiments of this type will be discussed below.

## 2.2. A Short Overview of Detection Techniques

There are several methods available for the detection of single nucleic acid molecules. Essentially, the most classical of these is electron microscopy. AFM and scanning near-field optical mi-

croscopy (SNOM) require less sophisticated sample preparation. In AFM, molecules can be used without labelling, while electron microscopy and SNOM require special preparation techniques. In spite of the advantage of very high resolution, all these methods have one disadvantage: their temporal resolution is low. This is the reason why conventional far-field microscopy has also found its firm place in single-molecule detection, and particularly in studies of single-molecule reactions. The temporal resolution of far-field microscopy with a simple CCD camera is 40 or 33 ms (PAL or NTSC systems, respectively) and can be increased to better than 1 ms with specialised cameras. It is no real problem to achieve single-molecule sensitivity similar to that of the competing techniques above. The apparent disadvantage of lower resolution is often not as severe as it may appear; in most cases single molecules are observed in the absence of other molecules in close vicinity. Then, the only price one has to pay is that the molecules appear to be too thick. Quite often this drawback is acceptable.

## 2.3. Conventional Far-Field Microscopy

In experiments where structural details are not required but the aim is to obtain information on reactions or the dynamics of a nucleic acid molecule, a conventional microscope is indeed the tool of choice. Mere visualisation of individual molecules is possible down to sizes of a few nanometres (see, for example, p. 274 of reference [7]). As indicated above, the price one has to pay is that the molecules appear too large, but for many applications this drawback is acceptable. The conventional fluorescence microscope has two significant advantages compared to the high-resolution near-field techniques described in Section 2.6: it is fast, that is, the dynamics of molecules and reactions become experimentally accessible, and preparation techniques can be easily adapted from bulk analytical methods.

One disadvantage of conventional far-field microscopy is that only molecules with a significant intrinsic fluorescence yield, or nonfluorescing molecules after extrinsic labelling with a fluorescent dye, can be investigated. Fortunately, DNA molecules can be easily labelled just by mixing them with suitable dyes such as DAPI, SybrGreen, YOYO, TOTO etc. The detection of single molecules by a fluorescence microscope is not so much a problem of the number of emitted photons. Since a molecule can, in principle, emit a light quantum every 10 ns, up to 100 million quanta might be expected from a single molecule. The problems are caused by other effects. First of all, a dye molecule will decay after a few hundred up to a million cycles of excitation and emission into a nonfluorescing derivative. This effect is usually termed "photobleaching". Secondly, to excite a dye molecule, some hundred million photons will pass through the detection volume and many of them may be scattered and reach the fluorescence detector, since, under these harsh conditions, optical filtering techniques are not completely sufficient to filter out all scattered excitation photons. Thus, it is difficult to identify the useful fluorescence photons. This is the major challenge for classical light microscopy

in single-molecule analysis, and not so much the sensitivity itself.

#### 2.4. Foerster/Fluorescence Resonance Energy Transfer (FRET)

Generally, fluorescence microscopy is by far the best-suited technique for observing single nucleic acid molecules in *living* cells. Other microscopic techniques, such as the electron microscope or the near-field microscopes described in Section 2.6, cannot look into their interior. Thus, approaches have been developed to circumvent the poor spatial resolution limit of conventional light microscopy. This is, to some extent, possible with Foerster (or fluorescence) resonance energy transfer (for a competent recent review, see reference [13]). FRET can increase the resolution beyond the diffraction limit of conventional microscopy. Here, a pair of fluorescent molecules is used to measure distances in the order of 1–10 nm. One molecule of this pair, the donor, absorbs light and usually re-emits it as fluorescence at a somewhat larger wavelength. If, however, the second molecule, the acceptor, with an absorption maximum close to the emission maximum of the donor, comes closer to the donor than the "Foerster" distance, the absorbed energy in the donor is transmitted to the acceptor which finally emits light at an even larger (more red) wavelength. Thus, a red shift indicates the mutual approaching of the donor/acceptor pair to a distance in the order of the Foerster distance, which is approximately 10 nm for the dye pair Cy3/Cy5, for example. These or other donor/acceptor pairs can be used to study protein–DNA interactions as well as RNA folding on a single-molecule basis. Generally, the Foerster distance  $R_0$  of a given donor/acceptor pair can be calculated as [Eq. (2)]:

$$R_0 = 8.79 \times 10^{-5} \kappa^2 n^{-4} \int \epsilon_A(\lambda) f_d(\lambda) \lambda^4 d\lambda \quad (2)$$

where the integral on the right is the overlap integral between the emission spectrum of the donor and the absorption spectrum of the acceptor. Only in rare cases can it really be calculated. However, one can estimate it graphically with good accuracy by comparing the overlap area for a desired donor/acceptor pair with the overlap area of a known pair such as Cy3/Cy5, which is available in datasheets of companies supplying these dyes.  $n$  is the refractive index of the environment of the dye molecules and  $\kappa$  is the "orientation factor".

This orientation factor is often not considered in the interpretation of FRET experiments. Optimal energy transfer is only achieved when the donor and acceptor molecules are in an ideal mutual position, rather like radio waves where the antenna has to be brought into optimal orientation with the radio emitter. If the orientation is nonoptimal, the orientation factor is too low and  $R_0$  turns out to be too small. When, in this case, the interaction of two molecules is studied by FRET, the result may be "no interaction" in spite of the fact that interaction does indeed occur. To be on the safe side, dyes should be attached to the molecules under investigation in such a way as

to enable them to rotate freely. Then  $\kappa^2$  is 2/3 and one can work with exact values.

The distance law for FRET is given by Equation (3):

$$E = 1/[1 + (R/R_0)^6] \quad (3)$$

It can be easily verified by Equation (3) that at a distance  $R = R_0$ , the energy-transfer efficiency  $E$  is just 1/2 (or 50%).

#### 2.5. Fluorescence Correlation Spectroscopy (FCS)

The basic principle of FCS is to extract information from a signal that appears, at first glance, just to be noise. This is the case if one illuminates a small volume and tries to detect photons at an angle of 90° with an extremely sensitive single-photon detector. The detector will produce signals due to thermal noise and to photons scattered statistically by the water. In such true noise, there is no correlation between a signal at a given time and a signal taken at a given later time. If, however, a single fluorescent molecule migrates sufficiently slowly through the observation volume, it will absorb and re-emit photons several times. As long as the fluorescent molecule is in the observation volume, the number of signals reaching the detector is slightly higher than that of the mere noise, since FCS is not a microscopic technique in the sense that images are evaluated. Rather, photon signals are detected statistically from a solution containing very few fluorescing molecules. Surprisingly, there is a strategy available that nevertheless gives the desired information. Measurements are made at a given time  $\tau$ , then after a delay time  $t$ ,  $2t$ ,  $3t$  and so on, the result for the different delay times is multiplied by that of  $t=0$  (that is, at time  $\tau$ ) and the results are summed up in a suitable way. This results in the autocorrelation function [Eq. (4)]:

$$G(\tau) = \langle F(\tau + t)F(t) \rangle_{\text{average over } t} / \langle F(t) \rangle_{\text{average over } t}^2 \quad (4)$$

where  $t$  is the waiting time. With a fast molecule, after a delay of, say,  $3t$ , there is no signal above average noise, since the molecule is no longer there, so the contribution of the corresponding mathematical product is zero, as it is for all longer delay times. With a slower molecule, there are also contributions for longer times, that is,  $G(\tau)$  becomes larger.

$G(\tau)$  is related to the speed of the molecule, which can be used to calculate the diffusion constant of the molecule and hence to estimate its molecular weight. Such experiments are often used to measure the binding of two individual molecules. This is possible when the two binding partners differ by at least a factor of 8–10 in size; for example, the binding of a small antigen to an antibody or the hybridisation of a small DNA or RNA oligonucleotide to a long DNA or RNA molecule. Maier et al.<sup>[14]</sup> studied binding sites of single glucocorticoid molecules on cell membranes. In such experiments, the smaller binding partner is made fluorescent so that it is visible as a fast-diffusing object, while the complex of small and large partners is visible as a slow object. The correlation analysis not only allows the absolute migrating velocities to be determined but also the ratio of free to bound molecules, that is, the bind-

ing constants according to the mass action law. The aforementioned size ratio of 8–10 is sometimes a drawback in FCS. However, this has been overcome by two-colour techniques,<sup>[15]</sup> occasionally even combined with FRET.<sup>[16]</sup>

## 2.6. Near-Field Microscopy (STM, SNOM, AFM)

As already indicated in Section 2.1, AFM can be used as a nanomanipulation instrument in the preparation of single DNA molecules, as well as a microscope with extremely high resolution. Essentially, however, the AFM is a near-field microscope. Near-field microscopes all work according to a common basic principle: a type of ultrafine stylus, the AFM tip, scans an object line by line and subsequently assembles an image from the data it has obtained. The near-field working principle can be clearly understood using the example of a doctor's stethoscope, where the acoustic waves of a beating heart with wavelengths of metres can be exploited to localise, for example, the human heart accurately within centimetres. The resolution limit of wave processes is no longer valid in the near field. Therefore, the spatial resolution of near-field microscopes can be much higher than the resolution of any far-field microscope. For the optical version of the near-field microscope, the scanning near-field optical microscope (SNOM) or optical stethoscope, the present practical resolution is 40–60 nm. For the scanning tunnelling microscope (STM), which uses quantum mechanical interactions to detect objects, it is under 1 nm and occasionally reaches subatomic resolution. Unfortunately, the STM can only image conducting molecules. In spite of the fact that DNA has a moderate conductivity (see also reference [80]), this is not sufficient to image DNA with optimal resolution. For this reason, early attempts to sequence DNA merely by imaging such a molecule with the extreme resolution of an STM were in the end unsuccessful.

Thus, the near-field microscope best suited to the analysis of single biomolecules is the AFM with a resolution in the low-nanometre range. Here, a combination of primarily very short-ranged van der Waals forces and electrostatic forces with a longer range is used for detection. To obtain the highest possible resolution, the electrostatic forces have to be suppressed. Even the combination of an AFM with a confocal laser scanning microscope is available.<sup>[17]</sup> To obtain an image of an object, one has to get as close to it as possible, since the resolution increases as the distance of the stylus from the molecule under investigation decreases. Voltage can easily be translated into position by using piezoelectric positioning elements, and an accuracy of a fraction of a nanometre is possible. Two detection modes are used: firstly, the distance of the stylus/tip to the molecule under investigation is readjusted at each scanning point, so that a preset force is acting. The position of the stylus allows the height of the object at this position to be determined. A second approach is to keep the *z* position of the stylus constant and to measure the force. Then, based on some assumptions about a specific distance law of this force, the height has to be calculated. By repeating this step pixel by pixel, finally the whole image can be assembled.

As discussed in the section on molecule preparation, the tip of an AFM can be used as an extremely fine manipulator for handling individual molecules. In addition, intramolecular forces in nucleic acids, for example during stretching, and intermolecular forces between nucleic acids and their binding partners can be measured. If one exerts a force of a few hundred piconewtons, noncovalent bonds will be cleaved. During pulling, the force first decreases approximately linearly with the pulling distance. Then, the force decreases abruptly to a much lower value. This effect is interpreted as the unbinding force. However, a number of other facts have to be considered, such as the pulling speed and specific molecular details. Nevertheless, the theory on single-molecule force spectroscopy has meanwhile been elaborated sufficiently to extract exact binding forces from such experiments. The unbinding process has a stochastic nature. Therefore, the measured unbinding forces of single unbinding events on the same individual molecule vary, and a statistical analysis has to be performed via a force histogram. This generates a probability curve with a most probable force of some tens up to a few hundred piconewtons and a full width at half maximum (FWHM) of 50–100 pN.<sup>[18,19]</sup> A large number of such force spectroscopy experiments have been performed with proteins. For example, a force-based protein biochip has been suggested<sup>[20]</sup> and different models of protein activation have been evaluated by single-molecule force spectroscopy.<sup>[21]</sup>

## 3. Results

### 3.1. Stretching and Folding of RNA and DNA

In textbook presentations the DNA molecule is always shown as an extended filamentous molecule, whereas in a physiological environment, that is, at salt concentrations around 150 mM, it is present as a globular molecule. It can be stretched by electrostatic forces,<sup>[45,22]</sup> with optical tweezers or with an AFM tip. With forces in the order of 10–20 pN, the DNA molecule can be extended into the normal structure with a length of about 0.3 nm per base pair. At forces of between 20 and 60 pN, the DNA molecule is reversibly overstretched up to 1.7-fold of its "natural" length<sup>[23,24]</sup> and is occasionally termed S-DNA. Above 60 pN, irreversible overstretching will occur. Similar effects are seen with *double*-stranded RNA. Using the tip of an AFM, a conformational change similar to the DNA B–S transition was detected, but overstretching is possible up to a factor of 2.<sup>[25]</sup> When small-molecule ligands bind to DNA, the force spectrum changes dramatically. For example, the binding of minor- or major-groove binders results in a transition from the conventional B form of DNA to its stretched S form, which manifests itself as a distinct plateau in the force–extension curve, while the intercalation of small molecules causes this plateau to vanish.<sup>[26]</sup>

The folding of single-stranded RNA, as opposed to double-stranded RNA, is different and can be better compared to protein folding. Indeed, some single-stranded RNA molecules fold into structures with enzymatic activity (ribozymes). Using single-molecule FRET, Russell et al.<sup>[27]</sup> showed that folding

occurs in deep valleys of the folding landscape, that is, along very stable pathways. Interestingly, the temporal sequence of folding is very critical. If long-range contacts are made at the right time, they stabilise a structure; if not, they may destabilise it. With similar techniques it was possible to show that the *Tetrahymena thermophila* ribozyme has eight different folding barriers. Forces of 10–30 pN are required to overcome these barriers.<sup>[28]</sup>

### 3.2. Chromatin

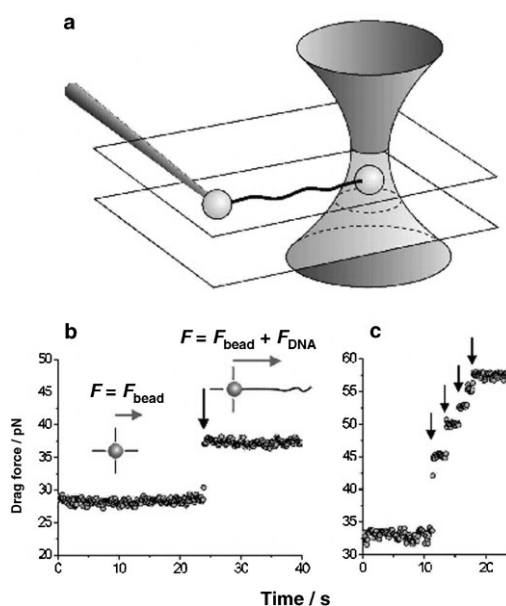
Of particular interest are the mechanical properties of chromatin, the complex of DNA with histones, which, in its most compact form, is known as the chromosome. Knowledge of the forces acting in chromatin and of the structural transitions of chromatin under the influence of a force is essential to understand the dynamics inside a cell nucleus. There, 2 m of DNA must be tightly packed to fit into a cell nucleus with linear dimensions of a few micrometers, but regions of chromatin also have to be unpacked to make the DNA available for transcription. It is reasonable to assume that strong forces are acting. Similar techniques as those mentioned above are used to study the stability of nucleosomes and of chromatin, the structures in which DNA is usually present in living cells. When a piece of chromatin consisting of DNA wound around nucleosomes is stretched, the first step is the release of 76 base pairs at low stretching force. When the force is increased, 80 additional base pairs are released in two stages.<sup>[29]</sup> An example of experiments of this type is given in Figure 2 (reproduced from reference [30]).

The assembly of nucleosomes and chromatin was studied in a reverse experiment,<sup>[31]</sup> and assembly was seen to be strongly inhibited at forces as low as 10 pN, which indicates that DNA needs its full flexibility to wrap around the nucleosomes. A short review on stretching of chromatin is available.<sup>[32]</sup>

### 3.3. Forces between Complementary DNA Molecules and in DNA Protein Complexes

Knowledge of the forces that govern the binding of a DNA strand to its complementary strand is not only interesting in principle, but also necessary to understand DNA melting, a process governing techniques such as DNA hybridisation, the polymerase chain reaction (PCR) or even DNA repair. In most corresponding experiments, one end of the molecule is attached to a surface. At its other end, one strand is also fixed to some point while the second strand is attached to an AFM tip or a bead that can be manipulated by magnetic or optical tweezers.

Stretched, unwound DNA was denatured (unzipped) locally with forces as low as 2 pN,<sup>[33]</sup> whereas 17–40 pN was needed to overcome the energy barrier for rupture of 9–13 kcal mol to completely separate short double helices.<sup>[34]</sup> Danilowicz et al. found multiple metastable intermediates when lambda phage DNA molecules were unzipped.<sup>[35]</sup> Unzipping changes progress along a DNA molecule in a sequence-dependent manner, but



**Figure 2.** Attaching a single  $\lambda$ -DNA molecule between two polystyrene beads. a) Schematic representation of a single DNA molecule suspended between two beads. After immobilisation of the first bead on a micropipette and the second in optical tweezers, end-biotinylated DNA is introduced into the flow cell. The deflection of the transmitted laser light through the polystyrene bead is used to monitor the drag force on the bead under a constant flow rate (approx.  $1000 \mu\text{m s}^{-1}$ ). b) The attachment of a single DNA molecule is detected as a sudden increase in drag force on the trapped bead. Under optimised conditions it is possible to have only a single molecule attached to the bead ( $0.25 \mu\text{g mL}^{-1}$   $\lambda$ -DNA). c) At a higher DNA concentration ( $2.5 \mu\text{g mL}^{-1}$ ) the drag force increases in a stepwise manner as a result of multiple DNA molecules attaching to the trapped bead. The vertical arrows indicate the increases in stepwise force detected per additional attached DNA molecule. (Reproduced from reference [30].)

not with sufficiently high specificity to be exploited for single-molecule DNA sequencing (see also below).

Binding between the two complementary strands of DNA is accomplished by stacking forces and by H bonds. In an experiment to study the effect of single base-pair (bp) mismatches, single-strand oligonucleotide molecules (20 bases in length) were bound to a surface, and their complementary strands (either ideally fitting or with one or two mismatches) were attached to the tip of an AFM.<sup>[36]</sup> Both types of strands were allowed to hybridise, then the rupture force was determined as a function of stacked bases or H bonds. Both types of bonds showed a linear relationship of approximately 4 pN per stacked base and 1.5 pN per H bond. Mismatches reduced the force by roughly this value, which indicates an internally consistent mechanism for hybridisation forces. Both the unzipping and re-joining of individual DNA strands have been observed on a single-molecule basis. A single nanopore made of  $\alpha$ -haemolysin is required. With this approach, the binding of an octanucleotide to its complementary DNA strand has been observed with techniques derived from patch clamping. In spite of the complex environment, the kinetics and binding constants for duplex formation were comparable to those of bulk experiments.<sup>[37]</sup>

Since about 15% of the 20000 human genes encode proteins that bind to nucleic acids, the forces which act during the binding of partner molecules of DNA such as proteins are almost as important as the DNA–DNA binding forces just described.

As an example of the binding of a protein with regulatory function, the binding of LexA, which controls 31 genes of *Escherichia coli* with slightly different binding motifs, to *E. coli* DNA were studied by force spectroscopy. Dissociation rates of 0.045 and 0.13 s<sup>-1</sup> were determined for the recA and yebG binding motifs, respectively. It was also possible to determine the width of the binding potentials as 5.4 and 4.9 Å, respectively. The most probable unbinding force was 40 pN.<sup>[19]</sup>

The binding of the regulatory ExpG protein to its target promoter regions in the soil bacterium *Sinorhizobium meliloti* 2011 was studied by a combination of molecular biology and force spectroscopy. Forces, dissociation rate constants and binding potential widths were determined with high accuracy and at different loading rates. Since the unbinding rate constants were different at low and high loading rates, it was possible to give some new details on the molecular binding mechanism.<sup>[38]</sup>

A new technique, unzipping force analysis of protein association (UFAPA),<sup>[39]</sup> detects DNA–protein interaction forces by unzipping a double-stranded DNA molecule with optical tweezers. When the unzipping fork reaches the site where a protein is bound, the unzipping force abruptly increases. Two classes of binding sites have been found for the DNA binding protein BsoBI: one with an average binding force of approximately 30 pN and one with a force of 45 pN.

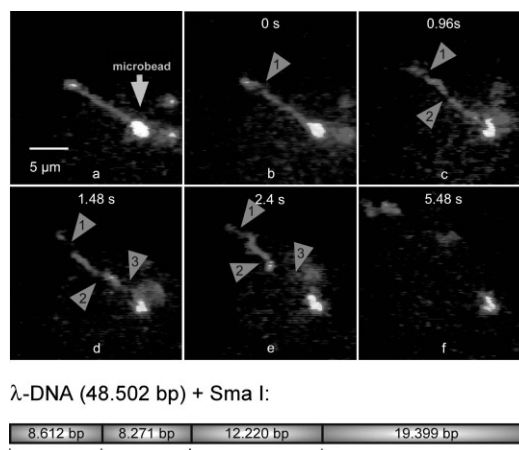
In the UFAPA experiment just described, a branch (the zippering front) was induced by exerting force. In vivo such a branch is, for example, the Holliday junction. Its migration represents an important step in genetic recombination and in DNA repair. It is generated by a protein complex called RudAB. Magnetic tweezers were used to measure forces and speeds during the migration of a Holliday junction along a 6-kb DNA molecule. It was found that, at zero force, the branch migrated at a rate of 43 bp s<sup>-1</sup>. It was directly demonstrated that RuvAB acts as a highly processive motor protein.<sup>[40]</sup>

### 3.4. Single-Molecule Restriction Analysis

UFAPA and similar techniques as described in the previous section can be used to identify and characterise single DNA molecules. Since a given kind of DNA binding protein attaches in a sequence-specific manner, sequence elements can be reconciled from such experiments. However, the present accuracy is approximately 25 bp<sup>[41]</sup> and thus only a relatively imprecise fingerprint can be obtained.

A much more precise technique is single-molecule restriction analysis. The enzymes involved in this technique, the restriction endonucleases, recognise sequence elements on long DNA molecules four to ten bases in length with single-base accuracy, and cut the molecule in a position close to this recognition sequence. This technique, in the bulk form known as DNA typing or paternity testing, has now also matured into a

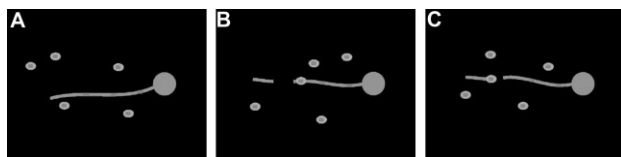
single-molecule technique. One successful approach in single-molecule DNA restriction analysis attaches the DNA molecule to be characterised to a micrometer-sized polystyrene microbead, which serves as a handle to be manipulated by optical tweezers.<sup>[42,43]</sup> The DNA molecules are labelled with the fluorescent dye CYBRGreen, which was previously found not to affect the reaction of the restriction endonucleases Apa1, Sma1 and EcoR1. The labelled DNA molecules were directly visible under a fluorescence microscope. When one of the restriction enzymes was added, the cutting of the DNA molecule could be observed directly. The restriction fragments floated away and could be sized just by measuring their fluorescence with an accuracy of approximately ±20%. Interestingly, for about 70% of all observed molecules the reaction progresses sequentially from one restriction site to the next on DNA molecules with more than one restriction site, for example in the reaction of Sma1 or EcoR1 with the lambda phage DNA molecule. Figure 3 shows microscopic images of such a reaction with Sma1.



**Figure 3.** a–f) Progress of the cutting of a single lambda phage molecule by a restriction endonuclease Sma I, which is expected to cut this molecule into four fragments according to the pattern at the bottom. For a similar result with EcoR1, which yields six fragments of lambda phage DNA, see reference [43].

Furthermore, two overall reaction rates were observed.<sup>[5,43]</sup> The reaction mechanism shown in Figure 4 was derived from such experiments. The model shows that two overall reaction velocities are expected. Indeed, for each restriction endonuclease studied so far, these two velocities were found.

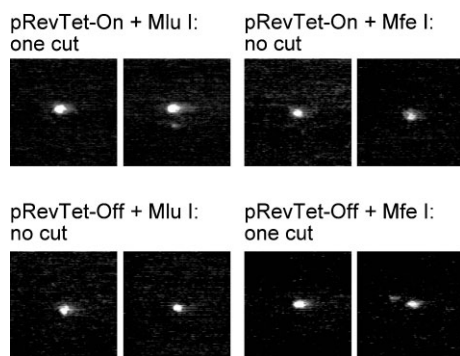
Meanwhile, techniques for sizing kilobase-sized DNA molecules have attained an accuracy of 7–14%.<sup>[44]</sup> Sometimes one would like to start the restriction reaction at a precisely selected time. For this purpose, use is made of the fact that most restriction endonucleases need free magnesium. If the latter is present as a caged compound, all ingredients for the restriction reaction can be premixed and the reaction then started by a short laser pulse, which liberates the magnesium and starts the reaction.<sup>[45]</sup> At sufficiently small concentrations of



**Figure 4.** A) An enzyme molecule is close to the DNA molecule attached to a polystyrene microbead and held in position with optical tweezers. Hydrodynamic flow stretches the DNA molecule. Shortly thereafter, the enzyme molecule attaches to the DNA and starts searching for a restriction sequence by linear diffusion along the DNA molecule. B) When the enzyme has found a restriction sequence, it cuts the DNA molecule. Here, it does not fall off along with the restriction fragment and can continue on the same molecule. This step is fast. C) In this case the enzyme molecule falls off and a new enzyme molecule has to find the DNA molecule. This step is slow.

caged calcium and with a sufficiently sharply focused laser, it is even possible to select the specific part of a long DNA molecule from which the reaction should proceed.<sup>[46]</sup>

For practical single-molecule DNA typing it is often not necessary to stretch the molecules. Restriction endonucleases also work on collapsed DNA molecules (see Figure 5).



**Figure 5.** Two different DNA molecules (the plasmids pRevTet-on and pRevTet-off) are treated with the restriction endonucleases MfeI and MluI. pRevTet-off is cut by MfeI (two fluorescence spots are visible) but not by MluI. For pRevTet-on the reverse is the case. As a consequence, both molecules are distinguishable in a comparably simple single-molecule experiment. (Figure courtesy of Bürk Schäfer, former PhD student of the IMB Jena.)

### 3.5. Single-Molecule DNA Hybridisation, Single-Molecule Sequencing and the Detection of DNA Damage on Single Molecules

A further strategy used to characterise DNA molecules is by DNA hybridisation. The logic behind it is the same as in restriction mapping, except that sequences in the DNA molecule to be characterised are not recognised by a protein, but by a complementary short sequence of DNA. To map megabase-sized individual DNA molecules,<sup>[47]</sup> they are stretched and attached to the surface of a microscope slide and short, kilobase-sized fluorescent DNA molecules are bound (hybridised) in a sequence-specific manner to the long target DNA. The resulting fluorescence pattern can be used to fingerprint the megabase-sized DNA molecule. This mapping can also be combined with two-photon imaging.<sup>[48]</sup> Again, a fingerprint-

like sequence-dependent pattern of fluorescence spots can be found.

DNA hybridisation is also the basis for single-molecule PCR reactions. Several experiments have shown that this reaction can be performed when only one DNA molecule is available. If such an individual DNA molecule is amplified and then sequenced, a very simple form of single-molecule DNA sequencing is obtained.<sup>[49]</sup> This is not really the technique termed "single-molecule sequencing" in the final decade of the last millennium. This expression meant a direct sequencing of one individual molecule without any amplification step. So far single-molecule sequencing has not lived up to its promises. However, optical sequencing of individual DNA molecules may receive a new impetus from work which enabled fluorochrome-labelled nucleotides to be incorporated exactly into an array of large single DNA molecules.<sup>[50]</sup>

No hybridisation step is required to detect abasic sites, that is, DNA damage where bases become lost in a DNA molecule. Such sites react with aldehyde reactive probes. If a streptavidin molecule is coupled to such a probe, the abasic sites can be detected with fluorochrome-coupled streptavidin, which tightly binds to the biotin label. Since such constructs reveal a very bright fluorescence, this sort of DNA damage can be easily detected on a single-molecule basis.<sup>[51]</sup>

Both abasic sites and DNA double-strand breaks can be observed on single molecules that are stained with the cyanine dye YOYO. Under intense illumination at wavelengths of 450–490 nm, T4 DNA molecules break and the breaks become directly visible. It has been shown that the flavonoid glycosyl hesperidin prevents double-strand breaks.<sup>[52]</sup> This observation is interesting for nutrition research since it is known that flavonoids in general prevent diseases, including cancers, which may be caused by DNA double-strand breaks. An ability to study the effect of flavonoids on DNA by direct inspection may significantly improve our knowledge of their protective effects.

### 3.6. Transcription Rates and Pausing of RNA Polymerases

Probably the best-investigated single-molecule enzyme reaction is that of RNA polymerases. These enzymes catalyse the synthesis of messenger RNA by reading the DNA template, that is, the transcription of genetic information into working molecules. Two questions are often asked: what is the speed of transcription and how uniform is it for a homogeneous preparation of RNA polymerase molecules? And what happens at pausing sites corresponding to regulatory elements in the DNA sequence to be read?

If the template DNA does not contain any sequence element that might cause pausing, the enzyme reads through the template at a relatively uniform speed. Some authors find a uniform kinetics among the individual molecules,<sup>[53]</sup> whereas others state inherent complexity.<sup>[54]</sup> One experiment studied 140 individual *E. coli* RNA polymerase molecules on a 2650-bp DNA template. The individual molecules were tagged with 199-nm-diameter microbeads. A Gaussian distribution of speeds was determined with a mean of 16 bp s<sup>-1</sup> and a standard deviation of 7 bp s<sup>-1</sup>, much larger than the accuracy of

the detection technique which was  $1 \text{ bps}^{-1}$ .<sup>[55]</sup> This result means that different individual RNA polymerase molecules do indeed work at different speeds.

As mentioned above, templates may contain sequence elements that cause pausing of the polymerase. The results reported in the literature are controversial. One work<sup>[56]</sup> reports that, at  $21^\circ\text{C}$  and a nucleotide concentration of  $1 \text{ mM}$ , the pauses are  $1\text{--}6 \text{ s}$  and their duration is not affected by either a hindering or an assisting load. On the contrary, it has been reported that, when the template DNA molecule is stretched, this pausing can be modified.<sup>[57]</sup> Also, by using runoff transcription to determine the position of the polymerase with an accuracy of  $5 \text{ bp}$ , the dwell time at the pausing sequences was reported to decrease with the assisting load<sup>[58]</sup> and to be sequence dependent.<sup>[59]</sup> Thus, the issue of the influences of load and sequence on transcriptional pausing is still open to discussion.

One of the aforementioned pausing sites may be a promoter. The complex sequence of events at a promoter site was studied for an RNA polymerase derived from the T7 bacteriophage. The binding of this enzyme was described by a  $k_{\text{off}}$  rate of  $2.9 \text{ s}^{-1}$ . The rate constant for the transition to the elongation mode was  $0.36 \text{ s}^{-1}$ . The reaction rate was  $43$  nucleotides per second and the transcript was released after approximately  $1200$  nucleotides.<sup>[60]</sup> With  $43$  nucleotides per second, this T7 bacteriophage polymerase is faster than the *E. coli* polymerase described above. It is also faster than lambda phage endonuclease, an enzyme performing essentially the reverse task of depolymerising a nucleic acid by removing base pairs, which works at a rate of  $12$  nucleotides per second,<sup>[61]</sup> comparable to the *E. coli* polymerase.

### 3.7. DNA Replication and DNA Repair

DNA replication and repair are performed by multienzyme complexes, such as DNA polymerases, helicases and topoisomerases. The working mechanism of some of these enzymes has also been elucidated in single-molecule studies.

UvrD is a helicase acting as a molecular motor that unwinds double-stranded DNA. It was concluded from bulk experiments that a UvrD molecule moves from the  $3'$  site to the  $5'$  site along one strand in a highly processive manner. Single-molecule experiments, however, revealed that the mechanism is more complex: after unwinding the double-stranded DNA molecule, the strands are unzipped in a step which is slow, since the enzyme has to be translocated. In contrast to the bulk experiments, single-molecule experiments have revealed that UvrD switches strands and translocates backwards on the other strand, probably to allow the DNA to re-anneal.<sup>[62]</sup>

Two other classes of enzymes that have been studied on a single-molecule basis are topoisomerases, enzymes which are involved in the unwinding of DNA,<sup>[63,64]</sup> and the enzymes of nucleotide excision repair (NER).<sup>[65]</sup> NER is a mechanism by which DNA damage on one of the two complementary strands of double-stranded DNA are repaired. One step in this DNA repair process is the recognition of the DNA damage by the XPA protein. The use of a laser-scanning microscope enabled

such damage in single DNA molecules to be visualised after fluorescence labelling of the DNA with Cy3.5. Their co-localisation with green fluorescent protein (GFP)-labelled XPA molecules could also be shown directly.

In the case of DNA polymerases, it has been shown that the rate of DNA replication slows down when the DNA molecule is stretched with a force of up to  $20 \text{ pN}$ . A higher force causes replication to cease completely.<sup>[66]</sup>

### 3.8. RNA, mRNP and Viruses in Single Cells

Inside a living cell, information stored in the DNA sequence and packed in the nucleus needs to be converted into working molecules such as proteins and to be trafficked to specific target sites. Single-molecule studies on *proteins* in living cells have been discussed in the reviews mentioned at the beginning of this overview.<sup>[3–5]</sup> Even more important are single-molecule studies on the behaviour of *nucleic acids* in single cells, since each individual nucleic acid molecule represents a message on its own. For example, a single messenger RNA molecule or a single mRNA–protein (mRNP) complex can be used, after migration from the cell nucleus into the cytoplasm, by the protein synthesis machinery (the ribosomes) to generate a large number of corresponding protein molecules.

Under optimal conditions, individual RNA molecules can be tracked for several hours. This has been shown in individual *E. coli* cells.<sup>[67]</sup> Most of the dynamics of such a molecule is consistent with Brownian motion, but occasionally a better description of the motion behaviour was achieved, when chain elongation was assumed. Implicitly this means that the transcription of an RNA molecule from the DNA template was observed on a single-molecule basis.

The observation that RNA moves essentially by Brownian motion or diffusion in *E. coli* cells may seem reasonable for these small cells without a nucleus, but this type of motion is less evident in eukaryotic cells. After transcription, the mRNA is packed together with proteins into a mRNP complex, it is then transported to a nuclear pore through which it is delivered to the cytoplasm, and finally to a ribosome where the message is translated into a protein molecule. There was some controversy as to whether this transport was accomplished by a directed transport process or by mere diffusion, as had been observed in bacterial cells. The controversy could not be settled by experiments based on bulk molecules. In a single-molecule experiment,<sup>[68]</sup> three positions of the transport process were labelled with blue, red and yellow fluorescent protein, all of which are variants of GFP. The genetic locus on the DNA was labelled with blue (cyan), the transcribed mRNA with red and the translated protein with yellow. Then the transport process was observed under a fluorescence microscope by single-particle tracking and with fluorescence recovery after photobleaching (FRAP), a technique which bleaches out a small volume in the cell and measures the motion of fluorescent molecules via the recovery of fluorescence in the bleached region. The result of this study was that mRNPs reach the nuclear pores and finally the site of translation in the cytoplasm by diffusion, with diffusion constants between  $0.01$  and  $0.09 \mu\text{m}^2 \text{ s}^{-1}$ . The process

is not directional. Only some steric hindrance due to corralling in the subnuclear fine structure slightly delays the process.

The opposite process, the path genetic information takes to get into a cell from outside, for example during a viral infection, was studied by using essentially the same technique<sup>[69]</sup> and observing viral RNPs with the genetic material of the influenza virus. Again, the major driving force behind the transport is diffusion through the cytoplasm to the nuclear pore. However, the interaction with the nuclear pore complex can have very different dissociation constants, which vary by two orders of magnitude. A specific protein, the M1 protein, downregulates the import of the influenza virus genome into the cell nucleus, but has no influence on other parts of the transport chain. It has been suggested that this effect can be exploited to develop drugs that will modify this interaction with the nuclear pore complex in such a way as to reduce the infectious potential of the influenza virus.

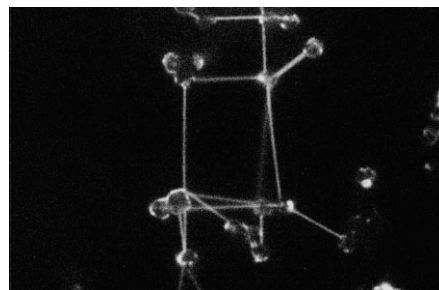
The transport of whole individual virus particles has been directly observed for individual adeno-associated virus particles by fluorescence microscopy. By comparing the reaction paths and speeds of many individual virus particles from the surface of the target cells into the nucleus, it has been possible to show that there are preferential reaction paths and that the average infection time is approximately 12 min, but with a large deviation among individual virus particles.<sup>[70]</sup>

### 3.9. Transistors, Nanomachines and Robots Made of Single DNA Molecules

Probably one of the most exciting perspectives, although in the early stages of feasibility studies, is nucleic acid nanotechnology.<sup>[71]</sup> This includes the use of DNA as a construction element in nanometre-sized structures such as transistors, computers or even automotive nanomachines. Two properties make DNA the ideal material for the assembly of complex nanostructures. First, a piece of a sequence from a single-stranded DNA molecule binds (hybridises) perfectly to its complementary sequence in a target molecule. For example, a sequence of four bases, of which  $4^4 = 256$  different versions are possible, has a length of 1.2 nm. Thus, there are a large number of addresses on a linearly extended DNA molecule that can, via intramolecular bonds, assemble into a predetermined structure and, via intermolecular bonds, form predetermined networks or allow binding of partner molecules with nanometre accuracy. For example, with the help of five 40-mer synthetic oligonucleotides, a 1669-mer DNA molecule folds by self hybridisation into a sharp geometrical octahedron.<sup>[72]</sup>

The partner molecule may be any other nano-object such as a carbon nanotube<sup>[73]</sup> or any other filament-like molecule tagged with a short sequence of DNA. Or it may itself be a DNA molecule, which can in turn bear a large number of addresses. Thus, one can easily envision the construction of nanostructures of an almost infinite complexity. The second interesting property of DNA results from the fact that the hybridisation of two complementary DNA strands depends on a number of physicochemical properties of the environment, such as temperature, pH value and the specific sequence itself.

By varying these properties it is possible to either enhance or, vice versa, to reduce hybridisation. Thus, not only static nanostructures, but also highly dynamic constructs of individual molecules are possible. Such nanostructures can be built even without the direct use of specific addresses. They can just be assembled with the help of optical tweezers (see Figure 6).



**Figure 6.** Rectangular nanostructure of DNA molecules assembled with optical tweezers. Each molecule bears a polystyrene microbead at its two ends by which the molecule can be manipulated with the tweezers. (Figure courtesy of L. Wollweber, IMB Jena.)

One disadvantage of approaches based on optical tweezers is that they are slow. Csaki et al. have shown how well the addressable single DNA or RNA molecules can be used as building blocks to construct molecular nanostructures via self-assembly.<sup>[74]</sup> Such nanostructures were used to assemble triangular patterns or rectangular grids of DNA molecules on mica surfaces. These structures were then covered using metal vapour deposition. Surprisingly, it was possible to peel off the whole metal layer from the DNA/mica substrate and thus obtain a negative print of the DNA pattern, that is, the DNA pattern had served as a template for printing the metal nanostructures. This approach, termed “molecular lithography”, uses sequence-directed self-assembly and is a bottom-up approach to generating nanostructures as opposed, for example, to photolithography which is a top-down approach.<sup>[75]</sup> The DNA molecules often have no intrinsic structures, while RNA molecules adopt a wide variety of functions and structures, such as triangles, spirals, rods and hairpins. A number of RNA molecules even reveal enzymatic activity (for an example, see reference [76]). In nucleic acid nanotechnology these RNA molecules can be used as building blocks (called tectosquares) to generate complex nanostructures. Such tectosquares have been used to construct, by mere self-assembly, arrays of hundreds of individual molecules with a wide variety of basic patterns.<sup>[77]</sup>

So far, not only complex structures have been constructed, but also functional devices have been assembled, for example field-effect transistors.<sup>[78]</sup> Building a field-effect transistor means precisely locating materials of different electric properties. An electron source and a drain have to be connected with a conductor. Single-walled carbon nanotubes are ideally suited for constructing a single-molecule transistor, but precise location is difficult. To achieve this with nanometre accuracy, a long DNA molecule is used as a scaffold. Shorter DNA sequences

that have bound molecules of a protein (in this case RecA, but other proteins would also be suitable) hybridise in a sequence-specific manner at exactly predetermined positions to the scaffold DNA molecule. The ends of the carbon nanotubes bear antibodies against RecA, that is, they recognise and bind exactly to the RecA blobs on the DNA scaffold. After incubating this assembly with an  $\text{AgNO}_3$  solution one finally ends up with the combination *source-carbon nanotube-drain*. Electronic testing proves that this assembly has the properties of a transistor, which suggests that DNA might have been used directly as an electric nanowire, since earlier experiments have found electric conductivity of DNA. This concept has, however, been challenged.<sup>[79]</sup> The study in question involved two experiments: the charge transfer from an electrode to a single DNA molecule was measured; then the dielectric response, when the molecule had no contact to any electrodes at its two ends, was determined. The authors found that DNA had a high resistivity, comparable to that of mica, glass or other electric insulators; that is, according to these findings, DNA is not suitable as an electric conductor. Meanwhile, it has been found that single DNA molecules can catch low-energy electrons,<sup>[80]</sup> so DNA is back again as a candidate for the construction of electronic devices, at least as a semiconducting molecule.

As indicated above, DNA can even be used to generate motion.<sup>[81]</sup> Since the two structural variants of DNA, the B and the Z forms, are of different length and thickness, the induction of a B-Z transition by changing, for example, the salt concentration may be used to induce motion in a DNA construction network. Knots have been tied with the DNA molecule<sup>[82]</sup> and DNA molecules have been suggested for use as molecular tweezers.<sup>[83]</sup> While the tweezers just open and close by moving their arms, a similar development results in a walking biped molecule, which has the perspective of becoming a transport molecule in single-molecule robotic devices.<sup>[84]</sup> The legs of the biped molecule consist of DNA with one given sequence for each leg. They are bound, by hybridisation, to another ladder-like DNA molecule acting as a sort of rail with the corresponding complementary sequences. When a third sequence is added which hybridises better to the rail than, say, the leading leg, the latter detaches but may re-bind to a position further ahead on the rail. Then the same process is repeated with the lagging leg. In the end the molecule has moved ahead along the rail DNA by one full step. Cyclic repetition of this process results in a "walking" molecule.

### 3.10. A Single-Molecule DNA Computer

Even a computer that performs its logical operations by single-molecule DNA hybridisation is possible. As in an electronic computer, numbers or other data elements are represented in a binary way by bits with the value 0 or 1. In the DNA computer described here,<sup>[85]</sup> the single-stranded hexanucleotide ATCACC corresponds, by definition, to 0 and GTCTGA corresponds to 1. A byte can be formed with a 24-mer nucleotide. For example, ATCACC ATCACC ATCACC ATCACC corresponds to 0000 (in the decadic number system used in everyday calculations this is 0), ATCACC ATCACC ATCACC GTCTGA to 0001

(the decadic 1), while ATCACC ATCACC GTCTGA ATCACC corresponds to 0010 (the decadic 2). To make the single oligonucleotide strands visible they are labelled with the fluorescence markers Cy5 (red) or Rhodamine Green. Calculations are performed by hybridisation, that is, the formation of double strands. The readout uses fluorescence correlation spectroscopy (see Section 2.5).

To calculate  $1 + 1 = 2$  (or  $0001 + 0001 = 0010$ ) a hybridisation system would have to be designed in which a double strand of two 0001 molecules would have a hybridisation behaviour identical to that of a single-strand 0010 molecule. This simple mathematical problem is a hard task for a DNA computer, which, in this field, will probably never become a serious competitor for electronic computers. For other logical operations, however, particularly when a large number of such operations have to be performed in parallel, the single-molecule DNA computer may become a true alternative. Such a task is the solution of the "satisfiability problem", which is described in detail in reference [85]. The mathematical problem is somewhat complex and it is beyond the scope of the present review on single-molecule DNA and RNA studies to describe it in detail. However, a quite simple mathematical problem can be solved with the experimental technique used by Schmidt et al.<sup>[85]</sup>: the question whether a given number is a prime number. For this purpose, a library of single-stranded oligonucleotides is synthesised and labelled with a red fluorescent dye containing all possible prime numbers ( $0001 = 1$ ,  $0010 = 2$ ,  $0011 = 3$ ,  $0110 = 5$ ,  $0111 = 7$ ,  $1011 = 11$ ,  $1110 = 13$ ). To determine whether a given number is a prime number, the complementary oligonucleotide with a green label is synthesised and added to the test library. The calculation is performed by hybridisation of the green molecules with the red molecules. Two outcomes are possible if the result is analysed by FCS. When hybridisation is detected, the tested number is a prime number. No hybridisation indicates that the tested number is not a prime number. Clearly, this problem would have been solved more easily with an electronic computer, even if a larger range of prime numbers were tested with longer oligonucleotides (up to 500 bases). However, detailed mathematical considerations indicate that there are indeed logical problems which require such a high degree of parallel computing that a single-molecule DNA computer might become the technique of choice.

## Summary and Outlook

DNA and RNA are the ideal molecules to be investigated and used on a single-molecule basis. Techniques for their manipulation and handling are now mature, so that subtle details of processes basic to life can be elucidated. Notably, DNA can be used equally well as a construction material for single-molecule devices such as robots or computers. It can be envisioned that single-molecule nucleic acid biotechnology will become a field of science on its own in the near future.

## Acknowledgments

This work was supported by the Stiftung Volkswagenwerk, grant I 75099 and by the German Research Ministry BMBF, grant 13N7506.

**Keywords:** DNA · microscopy · nanotechnology · RNA · single-molecule studies

- [1] B. Rothman, *Proc. Natl. Acad. Sci. USA* **1961**, *47*, 1981–1991.
- [2] Eds.: K. O. Greulich, E. Klose, H. Neumann, *Exp. Tech. Phys.* **1995**, *41*(2), 157–311.
- [3] K. O. Greulich, V. Uhl, *Encyclopedia of Life Sciences*, Macmillan, London, **2000**.
- [4] K. O. Greulich, *Curr. Pharm. Biotechnol.* **2004**, *5*, 243–259.
- [5] K. O. Greulich, *Encyclopedia of Molecular Biology and Medicine*, Vol. 9 (Ed.: M. Meyers), Wiley-VCH, Weinheim, **2005**, pp. 269–293.
- [6] Y. Ishii, T. Yanagida, *Single Mol.* **2000**, *1*, 5–13.
- [7] K. O. Greulich, *Micromanipulation by Light in Biology and Medicine: The Laser Microbeam and Optical Tweezers*, Birkhäuser, Basel, **1999**.
- [8] B. Nasanshargal, B. Schäfer, K. O. Greulich, in *Fluorescence Spectroscopy, Imaging and Probes* (Eds.: R. Kraayenhoff, A. J. W. G. Visser, H. C. Geritsen), Springer, Berlin, **2000**.
- [9] A. Bensimon, A. Simon, A. Chiffaudel, V. Croquette, F. Heslot, D. Bensimon, *Science* **1994**, *265*, 2096–2098.
- [10] A. Ashkin, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 4853–4860.
- [11] A. Hoffmann, G. Meyer zu Hörste, G. Pilarczyk, S. Monajembashi, V. Uhl, K. O. Greulich, *Appl. Phys. B* **2000**, *71*, 747–753.
- [12] T. R. Strick, V. Croquette, D. Bensimon, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 10579–10583.
- [13] R. M. Clegg, *BioPharm Int.* **2004**, *9*, 42–54.
- [14] C. Maier, D. Rünzler, L. Wagner, G. Grabner, G. Köhler, A. Luger, *Single Mol.* **2002**, *3*, 211–217.
- [15] T. Kohl, K. G. Heinze, R. Kuhlemann, A. Koltermann, P. Schwille, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12161–12166.
- [16] A. Amediek, E. Haustein, D. Scherfeld, P. Schwille, *Single Mol.* **2002**, *3*, 201–210.
- [17] M. Horton, G. Charras, C. Ballestrem, P. Lehenkari, *Single Mol.* **2000**, *1*, 135–138.
- [18] W. Grange, T. Strunz, I. Schumakovich, H. J. Güntherodt, M. Hegner, *Single Mol.* **2001**, *2*, 75–78.
- [19] F. Kühner, L. T. Costa, P. M. Bisch, S. Thalhammer, W. M. Heckl, H. E. Gaub, *Biophys. J.* **2004**, *87*, 2683–2690.
- [20] K. Blank, T. Mai, I. Gilbert, S. Schiffmann, J. Rankl, R. Zivin, C. Tackney, T. Nicolaus, K. Spinnler, F. Oesterheld, M. Benoit, H. Claussen-Schaumann, H. E. Gaub, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 11356–11360.
- [21] R. Nevo, V. Blumfeld, M. Elbaum, P. Hinterdorfer, Z. Reich, *Biophys. J.* **2004**, *87*, 2630–2634.
- [22] B. Schäfer, B. Nasanshargal, S. Monajembashi, H. Gemeinhardt, K. O. Greulich, V. Uhl, *Cytometry* **1999**, *36*, 209–216.
- [23] D. Anselmetti, J. Fritz, B. Smith, X. Fernandez-Busquets, *Single Mol.* **2000**, *1*, 53–58.
- [24] J. F. Marko, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 11770–11772.
- [25] M. Bonin, R. Zhu, Y. Klaue, J. Oberstrass, E. Oesterschulze, W. Nellen, *Nucleic Acids Res.* **2002**, *30*, e81.
- [26] R. Eckel, R. Ros, A. Ros, S. D. Wilking, N. Sewald, D. Anselmetti, *Biophys. J.* **2003**, *85*, 1968–1973.
- [27] R. Russell, X. Zhuang, H. P. Babcock, I. S. Millett, S. Doniach, S. Chu, D. Herschlag, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 155–160.
- [28] B. Onoa, S. Dumont, J. Liphardt, S. B. Smith, I. Tinoco, C. Bustamante, *Science* **2003**, *299*, 1892–1895.
- [29] B. B. Brower-Toland, C. L. Smith, R. C. Yeh, J. T. Lis, C. L. Peterson, M. D. Wang, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1960–1965.
- [30] M. L. Bennink, L. H. Pope, S. H. Leuba, B. G. de Groot, J. Greve, *Single Mol.* **2001**, *2*, 91–97.
- [31] G. Peterman, H. Sosa, L. S. B. Goldstein, W. E. Moerner, *Biophys. J.* **2001**, *81*, 2851–2863.
- [32] J. Zlatanova, S. H. Leuba, *J. Mol. Biol.* **2003**, *331*, 1–19.
- [33] T. R. Strick, V. Croquette, D. Bensimon, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 10579–10583.
- [34] L. H. Pope, M. C. Davies, C. A. Laughton, C. J. Roberts, S. J. B. Tendler, P. M. Williams, *Eur. Biophys. J.* **2001**, *30*, 53–62.
- [35] C. Danilowicz, V. W. Coljee, C. Bouzigues, D. K. Lubensky, D. R. Nelson, M. Prentiss, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1694–1699.
- [36] B. D. Sattin, A. E. Pelling, M. C. Groh, *Nucleic Acids Res.* **2004**, *32*, 4876–4883.
- [37] S. Howorka, L. Movileanu, O. Braha, H. Bayley, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12996–13001.
- [38] F. W. Bartels, B. Baumgarth, D. Anselmetti, R. Ros, A. Becker, *J. Struct. Biol.* **2003**, *143*, 145–152.
- [39] S. Koch, M. D. Wan, *Phys. Rev. Lett.* **2003**, *91*, 28103.
- [40] A. Dawid, V. Croquette, M. Grigoriev, F. Heslot, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11611–11616.
- [41] S. J. Koch, A. Shundrovsky, B. C. Jantzen, M. D. Wang, *Biophys. J.* **2002**, *83*, 1098–1105.
- [42] B. Schäfer, H. Gemeinhardt, V. Uhl, K. O. Greulich, *Single Mol.* **2000**, *1*, 33–40.
- [43] B. Schäfer, H. Gemeinhardt, K. O. Greulich, *Angew. Chem.* **2001**, *113*, 4799–4802; *Angew. Chem. Int. Ed.* **2001**, *40*, 4663–4666.
- [44] S. Laib, M. Rankl, T. Ruckstuhl, S. Seeger, *Nucleic Acids Res.* **2003**, *31*, 138–143.
- [45] C. Hoyer, S. Monajembashi, K. O. Greulich, *J. Biotechnol.* **1996**, *52*, 65–73.
- [46] V. Namasivayan, R. G. Larson, D. T. Burke, M. A. Burns, *Anal. Chem.* **2003**, *75*, 4188–4194.
- [47] Y. K. Wang, E. Huff, D. C. Schwartz, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 165–169.
- [48] K. König, A. Göhlert, T. Liehr, I. F. Loncarevic, I. Riemann, *Single Mol.* **2000**, *1*, 41–52.
- [49] M. B. Wabuyele, S. A. Soper, *Single Mol.* **2001**, *2*, 13–22.
- [50] A. Ramanathan, E. J. Huff, C. C. Lamers, K. D. Potamouisis, D. K. Forrest, D. C. Schwartz, *Anal. Biochem.* **2004**, *330*, 227–241.
- [51] T. Hirose, T. Ohtani, H. Muramatsu, A. Tanaka, *Photochem. Photobiol.* **2002**, *76*, 123–126.
- [52] Y. Yoshikawa, M. Suzuki, N. Yamada, K. Yoshikawa, *FEBS Lett.* **2004**, *566*, 39–42.
- [53] K. Adelman, A. La Porta, T. J. Santangelo, J. T. Lis, J. W. Roberts, M. D. Wang, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 13538–13543.
- [54] M. Guthold, D. A. Erie, *ChemBioChem* **2001**, *2*, 167–170.
- [55] S. F. Tolic-Norrelykke, A. M. Engh, R. Landick, J. Gelles, *J. Biol. Chem.* **2004**, *278*, 3292–3299.
- [56] K. C. Neumann, E. A. Abbondanzieri, R. Landick, J. Gelles, S. M. Block, *Cell* **2003**, *115*, 437–447.
- [57] N. R. Forde, D. Izhaky, G. R. Woodcock, G. J. L. Wuite, C. Bustamante, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11682–11687.
- [58] A. Shundrovski, T. Santangelo, J. W. Roberts, M. D. Wang, *Biophys. J.* **2004**, *87*, 3945–3963.
- [59] L. Bal, A. Shundrovsky, M. D. Wang, *J. Mol. Biol.* **2004**, *344*, 335–349.
- [60] G. M. Skinner, C. G. Baumann, D. M. Quinn, J. E. Molloy, J. G. Hoggatt, *J. Biol. Chem.* **2004**, *279*, 3239–3244.
- [61] T. T. Perkins, R. V. Dalal, P. G. Mitsis, S. M. Block, *Science* **2003**, *301*, 1914–1921.
- [62] M. N. Designes, T. Lionnet, X. G. Xi, D. Bensimon, V. Croquette, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6439–6444.
- [63] T. R. Strick, V. Croquette, D. Bensimon, *Nature* **2000**, *404*, 901–904.
- [64] N. H. Dekker, V. V. Rybenkov, M. Duguet, N. J. Crisona, N. R. Cozzarelli, D. Bensimon, V. Croquette, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12126–12131.
- [65] G. M. J. Segers-Nolten, C. Wyman, N. Wijgers, W. Vermeulen, A. T. M. Lenferink, J. H. G. Hoeijmakers, J. Greve, C. Otto, *Nucleic Acids Res.* **2002**, *30*, 4720–4727.
- [66] A. Goel, M. D. Frank-Kamenetskii, T. Ellenberger, D. Herschbach, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8485–8489.
- [67] I. Golding, E. C. Cox, *Proc. Natl. Acad. Sci. USA* **2001**, *101*, 11310–11315.
- [68] Y. Shav-Tal, X. Darzacq, S. M. Shenoy, D. Fusco, S. M. Janicki, D. L. Spectator, R. H. Singer, *Science* **2004**, *304*, 1797–1800.
- [69] H. P. Babcock, C. Chen, X. Zhuang, *Biophys. J.* **2004**, *87*, 2749–2758.
- [70] G. Seisenberger, M. U. Ried, T. Endreß, H. Büning, M. Hallek, C. Bräuchle, *Science* **2001**, *294*, 1929–1932.

- [71] H. Yan, *Science* **2004**, *306*, 2048–2049.
- [72] W. M. Shih, J. D. Quispe, G. F. Joyce, *Nature* **2004**, *427*, 618–622.
- [73] M. Zheng, A. Jagota, M. S. Strano, A. P. Santos, P. Barone, G. Chou, B. A. Diner, M. S. Dresselhaus, R. S. Mclean, G. B. Onoa, G. G. Samsonidze, E. D. Semke, M. Usrey, D. J. Walls, *Science* **2003**, *302*, 1545–1548.
- [74] A. Csaki, G. Maubach, D. Born, J. Reichert, W. Fritzsche, *Single Mol.* **2002**, *3*, 275–280.
- [75] Z. Deng, C. Mao, *Angew. Chem.* **2004**, *116*, 4160–4162; *Angew. Chem. Int. Ed.* **2004**, *43*, 4068–4070.
- [76] D. Rueda, G. Bokinsky, M. M. Rhodes, M. J. Rust, X. Zhuang, N. G. Walter, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10066–10071.
- [77] A. Chworos, I. Severcan, A. Y. Koyfman, P. Weinkam, E. Oroudjev, H. G. Hansma, L. Jaeger, *Science* **2004**, *306*, 2068–2072.
- [78] K. Keren, R. S. Berman, E. Buchstab, U. Sivan, E. Braun, *Science* **2003**, *302*, 1380–1384.
- [79] G. Gomez-Navarro, F. G. Moreno-Herrero, J. de Pablo, J. Colchero, C. Gomez-Herrero, A. M. Baro, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8484–8487.
- [80] S. G. Ray, S. S. Daube, R. Naaman, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15–19.
- [81] A. Mao, W. Sun, Z. Shen, N. C. Seeman, *Nature* **1999**, *397*, 144–146.
- [82] T. Arai, R. Yasuda, K. I. Akashi, Y. Harada, H. Miyata, K. Kinoshita, H. Ito, *Nature* **1999**, *399*, 446–448.
- [83] N. D. Rowell, *BioPharm Int.* **2000**, *9/10*, 86–87.
- [84] W. B. Sherman, N. C. Seeman, *Nano Lett.* **2004**, *4*, 1203–1207.
- [85] K. A. Schmidt, C. V. Henkel, G. Rozenberg, H. P. Spaink, *Nucleic Acids Res.* **2004**, *32*, 4962–4968.

---

Received: January 20, 2005

Revised: May 26, 2005