

High-resolution atomic force microscopy of duplex and triplex DNA molecules

Dmitry Klinov^{1,2}, Benjamin Dwir¹, Eli Kapon¹, Natalia Borovok³,
Tatiana Molotsky³ and Alexander Kotlyar^{3,4}

¹ Laboratory of Physics of Nanostructures, Ecole Polytechnique Fédérale de Lausanne (EPFL), Station 3, CH-1015 Lausanne, Switzerland

² Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia

³ Department of Biochemistry, George S Wise Faculty of Life Sciences, Tel-Aviv University, Israel

⁴ Institute of Nanoscience, Tel-Aviv University, Israel

E-mail: benjamin.dwir@epfl.ch and s2shak@post.tau.ac.il (Alexander Kotlyar)

Received 25 January 2007, in final form 22 March 2007

Published 4 May 2007

Online at stacks.iop.org/Nano/18/225102

Abstract

Double-stranded poly(dG)–poly(dC) and triple-stranded poly(dG)–poly(dG)–poly(dC) DNA were deposited on the modified surface of highly oriented pyrolytic graphite (HOPG) and visualized using atomic force microscopy with high-resolution (radius of ~ 1 nm) tips. The high resolution attained by this technique enabled us to detect single-stranded regions in double-stranded poly(dG)–poly(dC) and double-stranded and single-stranded regions in poly(dG)–poly(dG)–poly(dC) triplexes, as well as to resolve the helical pitch of the triplex molecules. We could also follow the reaction of G-strand extension in poly(dG)–poly(dC) by the Klenow exo^- fragment of DNA polymerase I. This approach to molecular visualization could serve as a useful tool for the investigation of irregular structures in canonical DNA and other biopolymers, as well as studies of the molecular mechanisms of DNA replication and transcription.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Investigation of the structure of single DNA molecules and of biochemical processes on a single-molecular level is essential for understanding the mechanisms of DNA replication and transcription. One of the most powerful tools for imaging the structure of single macromolecules is the atomic force microscope (AFM) [1]. It has been proven to be very useful for the investigation of various nanostructures such as carbon nanotubes [2], proteins [3] and other polymers [4].

However, imaging by AFM requires deposition of the molecules on surfaces, and the interaction with the surface strongly affects the geometry and the properties of ‘soft’ biomolecules [5, 6]. For example, the height of double-stranded DNA molecules adsorbed on mica (the most common substrate for AFM studies of DNA and other biomolecules) and imaged by AFM (0.6–0.8 nm) [7] is much lower than

the diameter of the molecules measured by nuclear magnetic resonance (NMR) in solution (1.7–1.9 nm) [8, 9] or by x-ray diffraction (XRD) in crystals (2 nm) [10, 11]. This reduction in DNA height is probably due to compression of the molecule and distortion of its internal helical structure as a result of its interaction with the surface. The level of deformation depends on the nature of the DNA and on the surface used as a substrate for the molecules [6]. There are very few studies of DNA using alternative surfaces, like atomically flat graphite [12] and gold [13], mainly due to the absence of efficient methods of DNA deposition on these surfaces.

Another problem associated with AFM imaging of biomolecules is the need for high resolution in order to image such small (< 2 nm) objects, especially when the internal structure of the molecules is of interest. The relatively large apex radius of commercially available Si AFM probes (5–10 nm) allows neither the determination of the internal

molecular features of DNA nor the detection of unusual folding regions in DNA resulting from either base mismatches or non-canonical base pairing.

In order to observe morphological features and monitor structural alterations in DNA we have developed a novel high-resolution AFM imaging technique that enables detailed characterization of DNA morphology, both in width and height. We have used chemically modified highly oriented pyrolytic graphite (HOPG) to adsorb DNA [14], high-resolution AFM probes and low-amplitude (3–5 nm) tapping mode imaging to reduce tip–molecule interactions. We performed detailed height and width AFM analysis of two types of molecules, poly(dG)–poly(dC) and poly(dG)–poly(dG)–poly(dC), composed of 1000 base pairs and 1000 triads, respectively. Both types of DNA were synthesized by the Klenow *exo*[−] fragment of DNA polymerase I as described in our recent papers [15, 16]. The high-resolution AFM imaging technique enabled us to identify irregularities in the structures: single-stranded loops along poly(dG)–poly(dC) regions, and single- and double-stranded loops and tails along triple-stranded DNA. Moreover, by identifying double- and triple-stranded regions along the synthesized poly(dG)–poly(dG)–poly(dC) molecules we could monitor the dynamics of the triplex synthesis on a molecular level, demonstrating its continuous, linear nature.

2. Materials and methods

2.1. Materials

Unless otherwise stated, reagents were obtained from Sigma-Aldrich (USA) and were used without further purification. 2'-deoxyribonucleoside 5'-triphosphates were purchased from Sigma-Aldrich (USA). Klenow fragment exonuclease minus of DNA polymerase I from *Escherichia coli* lacking the 3' → 5' exonuclease activity (Klenow *exo*[−]) was purchased from Fermentas (Lithuania). The oligonucleotides were purchased from Alpha DNA (Montreal, Canada). 1000 bp poly(dG)–poly(dC) was synthesized as described in our recent publication [15].

2.2. DNA polymerase assays

DNA triplex synthesis was conducted essentially as described in our recent paper [16]. The reaction was conducted at 37 °C in a standard assay mixture containing 60 mM potassium phosphate buffer (KPi) (pH 7.4), 5 mM dithiothreitol (DTT), 1 mM MgCl₂, 0.5 mM dGTP and 0.25 μM (in molecules) 1000 base pairs poly(dG)–poly(dG) at 37 °C. The G-strand extension was started by adding 0.2 μM Klenow *exo*[−] fragment of DNA polymerase I.

2.3. High-performance liquid chromatography (HPLC) separation of the polymerase products

The synthesis was halted by adding 5 mM EDTA to the assay. The aliquots were withdrawn from the assay and passed through size-exclusion HPLC in order to separate high-molecular-weight products of the synthesis from nucleotides MgCl₂, and other reaction components of the assay. The

separation was carried out with a TSK-gel G-DNA-PW HPLC column (7.8 mm × 300 mm) from TosoHaas (Japan) by isocratic elution with 20 mM Tris-acetate, pH 7.5, at a flow rate of 0.5 ml min^{−1}. The injection volumes were 40–150 μl. All experiments were conducted on an Agilent 1100 HPLC system, with a photodiode array detector. Peaks were identified from their retention times obtained from the absorbance at 260 nm. The peak corresponding to a high-molecular-weight product of the synthesis was collected and used for deposition on modified HOPG.

2.4. Deposition of DNA on modified HOPG

We deposited DNA on modified HOPG surfaces essentially as described in our recent paper [14]. Freshly cleaved HOPG (NT-MDT, Russia) surfaces were incubated with the graphite modifier (GM) solution (Nanotuning, Chernogolovka, Russia, <http://www.nanotuning.com>) for 15 min. The graphite modifier is composed of a hydrocarbon–peptide pair terminating with an amine group: (CH₂)_n(NCRH₂CO)_m–NH₂. The application of this modifier to the hydrophobic HOPG surface results in the formation of a thin (0.7 nm) monolayer of positively charged polymer on the graphite surface, thus promoting DNA binding. The surface roughness of the modified HOPG is approximately 0.1–0.2 nm, thus high-resolution AFM imaging is possible.

The DNA fraction eluted from the HPLC column was diluted 10-times in 2 mM Tris-acetate buffer (pH 7.0); 10 μl of the DNA solution (concentration in molecules is approximately equal to 1–2 nM) was applied onto the modified HOPG surface and incubated for 10 min at room temperature. The DNA solution was then removed from the surface by letting N₂ gas flow from the centre to the edge of the substrate.

2.5. Atomic force microscopy (AFM)

The samples were imaged by Nanoscope III (Veeco, USA) and Ntegra (NT-MDT, Russia) AFMs in the tapping mode. The free oscillating amplitude of the AFM cantilever was around 3–4 nm; imaging was thus performed in 'light tapping' mode (for which the set-point amplitude is close—0.98%—to the free-oscillating amplitude of 4 nm). High-resolution AFM tips were prepared by plasma-assisted growth of carbon spikes at the apex of Olympus AC160TS Si probes [17] with force constant 42 N m^{−1} and resonant frequency 300 kHz. This growth did not change the force/frequency parameters of the cantilevers. The quality of the tips was controlled with an EM 100CX transmission electron microscope (JEOL, Japan) under an accelerating voltage of 80 kV (see figure 1). In spite of the multiplicity of tips produced by this method, there is only one that protrudes farther than the others (by ~100 nm) and is the one that contacts the sample, thus producing the high-resolution image (see figure 1 inset). This is made possible by the small roughness (a few nanometres at most) of the substrate and the small height of the imaged molecules, which are both much smaller than the distance between the imaging tip and the 'secondary tips'. In most cases, height image was used; however the tip oscillation amplitude and phase images were stored as well (on the Ntegra AFM) and used when needed. All AFM images were 'flattened' (the average second-order polynomial fit was calculated and subtracted from the image

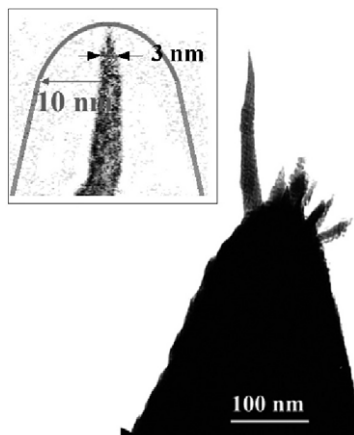


Figure 1. Transmission electron microscope image of the high-resolution AFM tip (length ~ 200 nm, apex radius ~ 1 nm). The inset is a magnified image of the end of the most protruding tip, demonstrating its small effective diameter as compared to the diameter of a typical Si commercial tip (radius = 10 nm, continuous line).

data for each image line), then analysed by the respective AFM's software and by the SPIP analysis software (Image Metrology, Denmark). In many cases, hundreds of cross-sections of many molecules were taken and the height/width of the molecule measured manually by SPIP. These values were then analysed statistically to obtain average/SD values, using about 100 measurements per data point, especially for figures 3 and 5.

3. Results and discussion

The high-resolution method of molecular visualization employed here makes it possible to detect various structural motifs along a single DNA molecule. This was demonstrated in several studies of DNA derivatives (poly(dG)–poly(dC) and poly(dG)–poly(dG)–poly(dC)) deposited on modified HOPG surfaces as we show below.

3.1. Structure of triplex DNA molecules

We performed detailed morphological studies on poly(dG)–poly(dG)–poly(dC) triplex DNA by high-resolution AFM. The molecules were synthesized by Klenow exo– using a recently published method [16]. The molecules were deposited on modified HOPG and imaged by a high-resolution tip (radius about 1 nm, estimated from the TEM image) in ‘light tapping’ mode. As seen in figure 2, three different structural motifs, with heights of 0.4 ± 0.1 nm, 1.0 ± 0.1 nm and 2.0 ± 0.1 nm, respectively, can clearly be distinguished along the molecules. In our recent study [14] we have shown that the average apparent height of single-stranded (ss) DNA, double-stranded (ds) DNA and triplex DNA deposited on the modified HOPG surface is about 50% higher than on a mica surface, and for triplex DNA the apparent height is similar to the real DNA diameter as measured in solution. This demonstrates that the structure of DNA molecules is less affected by the modified HOPG surface as compared to mica, thus less ‘flattening’ of the DNA is expected. It should also be noted that, when imaged with the conventional AFM technique (deposition on mica and scanning with commercial tips), no clear distinction

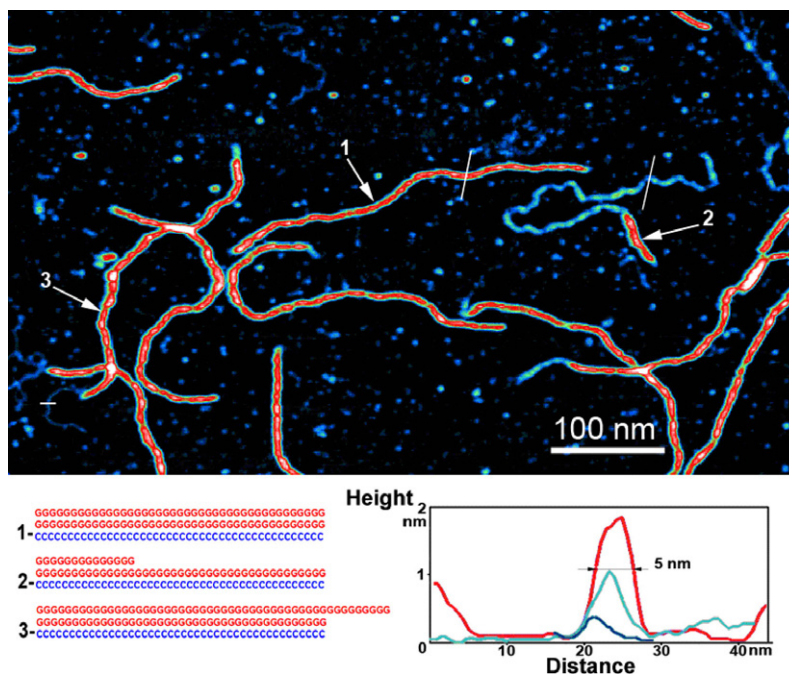


Figure 2. High-resolution AFM image of poly(dG)–poly(dG)–poly(dC) triplex DNA. Notice the single-strand (dark blue), double-strand (light blue) and triple-strand (red) regions in the molecules. The bottom-right inset is a cross-section of triplex, dsDNA and ssDNA regions in the molecule, taken at the short unmarked lines in the image. Bottom-left diagrams present schematic views of three types of molecules, as indicated by the arrows: 1, full-length or ‘completed’ triplex molecule; 2, ‘partially-synthesized’ triplex, containing both triplex and dsDNA fragments; 3, ‘over-synthesized’ triplex, containing a single-stranded fragment at the end of the molecule.

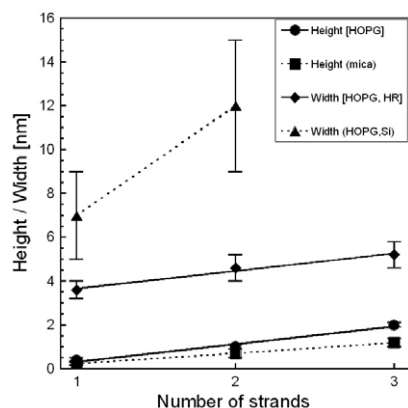


Figure 3. Dependence of the height (bottom traces) and the width (top traces) of molecules on the number of strands composing the DNA. Bottom trace (squares): DNA height on mica (from [14]). Circles: DNA height on modified HOPG. Diamonds: DNA width on HOPG, using high-resolution AFM tip. Top trace (triangles): DNA width on modified HOPG, using conventional AFM tip (from [18]). The observed width is the result of a geometrical convolution of the real molecular width and the tip radius. The fitting (linear least-square) lines of the DNA height and width (measured with high-resolution probes) have the same slope and their difference is 3.3 nm.

of the structural alterations along the triplex molecules could be obtained. This attests to the higher spatial resolution of the AFM approach used here. We suggest that the three motifs detected by the high-resolution AFM (figure 2) correspond to single-, double- and triple-stranded DNA. This is supported by previous conventional AFM measurements of ssDNA, standard double-stranded plasmid DNA, poly(dG)–poly(dC) molecules and triplex DNA on modified HOPG, where the height of each type of DNA has been measured [14]. The height data thus help us to identify the type of DNA segment in the image (see inset in figure 2). The dependences of the height and the width of different parts of the DNA molecules (averaged from hundreds of cross-sections similar to those shown in figure 2, see section 2.5) on the number of strands of which they are comprised are presented in figure 3. As seen in this figure, both the molecular height and the width depend linearly on the number of strands composing the DNA. The measured height on HOPG is higher than the height on mica, as we have already demonstrated [14]. The apparent width, measured by high-resolution probes on HOPG, is consistently higher than the molecular height, as shown by the two full lines in figure 3 having the same slope. This constant difference (3.3 nm) is at least partially due to the convolution of the real molecule's width with the tip radius, for which we have an estimated value of 1 nm (from the TEM image of the tip, figure 1). Therefore it is also plausible that the molecule bound to the surface does not present a round cross-section to the AFM; its deconvoluted width is about 1–1.5 nm greater than its height. Similar AFM images of ssDNA and dsDNA done on HOPG using conventional Si probes have yielded much higher values [18], as shown in figure 3 (triangles and top dotted line). It should be noted that the small radius of our tips allows such small values of DNA width to be measured, thus allowing us to visualize morphological features of the molecules, as will be shown below (section 3.4).

The molecule denoted by number 2 in figure 2 illustrates another interesting feature: the triplex part is much stiffer than the double-stranded part of the molecule. This is proven by the fact that under identical deposition and measurement conditions, triplex parts look higher in AFM than dsDNA parts, although they have the same diameter in liquid. We calculated the ratio of end-to-end distance to contour length for 1000 base pair (or triad) long DNA: pSK+ plasmid dsDNA, poly(dG)–poly(dC) dsDNA and poly(dG)–poly(dG)–poly(dC) triplex, all deposited on HOPG. This ratio, which measures the stiffness of the molecules ($1 = \text{rod-like molecule}$) was 0.4 ± 0.2 for 1000 bp long section of pSK+, 0.3 ± 0.2 for 1000 bp long poly(dG)–poly(dC) and 0.8 ± 0.3 for 1000 triad long triplex, indicating its higher stiffness. We also calculated the persistence length of the triplex on HOPG, which is 140 ± 10 nm (average of a few tens of molecules). It is somewhat higher than that of dsDNA (about 50 nm [19]), pointing to the increased stiffness of the triplex molecules.

3.2. AFM visualization of the synthesis process of triplex DNA

The poly(dG)–poly(dG)–poly(dC) molecules have been produced by enzymatic extension of the G-strand in poly(dG)–poly(dC) [16]. Essentially, the synthesis proceeds by a sequential elongation of the poly(dG) strand of the original double-stranded poly(dG)–poly(dC) molecule, and folding of polyG *de novo* back to form the triplex. In view of this understanding, the image in figure 2 clearly depicts different stages in the synthesis: we can distinguish 'completed' triplex molecules, 'non-completed' ones (consisting of both triple- and double-stranded fragments) and some 'over-synthesized' molecules, containing a short poly(dG) overhang at the end of the triplex molecule. The presence of both non-completed and over-synthesized triplexes in the same enzymatic assay is probably related to the statistical distribution of the rates of dG-strand elongation in different polyG–polyC molecules by the enzyme.

To obtain more insight into the synthesis process, we imaged successive stages in the reaction by sampling the assay after different durations, from 0 (before the addition of the enzyme) to 6 h. At each time the sampled molecules were deposited on the modified surface of HOPG and visualized by AFM using high-resolution tips (see figure 4). In this way we could obtain a series of 'snapshots' of successive stages within the synthesis process. As seen in figures 4(b)–(f), the length of the triplex fragment (seen as a bright fragment in the figure) increases with the synthesis time, while the length of the remaining double-stranded fragment (seen as a narrower and darker fragment in the figure) decreases. As before, the difference in the brightness in the image of the fragments along the synthesized molecule is due to the difference in the height between the triplex and the dsDNA, 2 ± 0.1 nm and 1.0 ± 0.1 nm, respectively [14]. The dependence of the relative average length of the triplex fragment on the duration of the synthesis is plotted in figure 5. The error bars represent the spread in the distribution of the length values of the triplex fragment at each step of the synthesis. It is between 15% and 20% for all data points. From the graph it is seen that the length of the triplex fragment increases with the synthesis time as a first-order process, with a time constant of 1.5 ± 0.16 h. These data are consistent with the synthesis data in [16], especially the FRET data provided there.

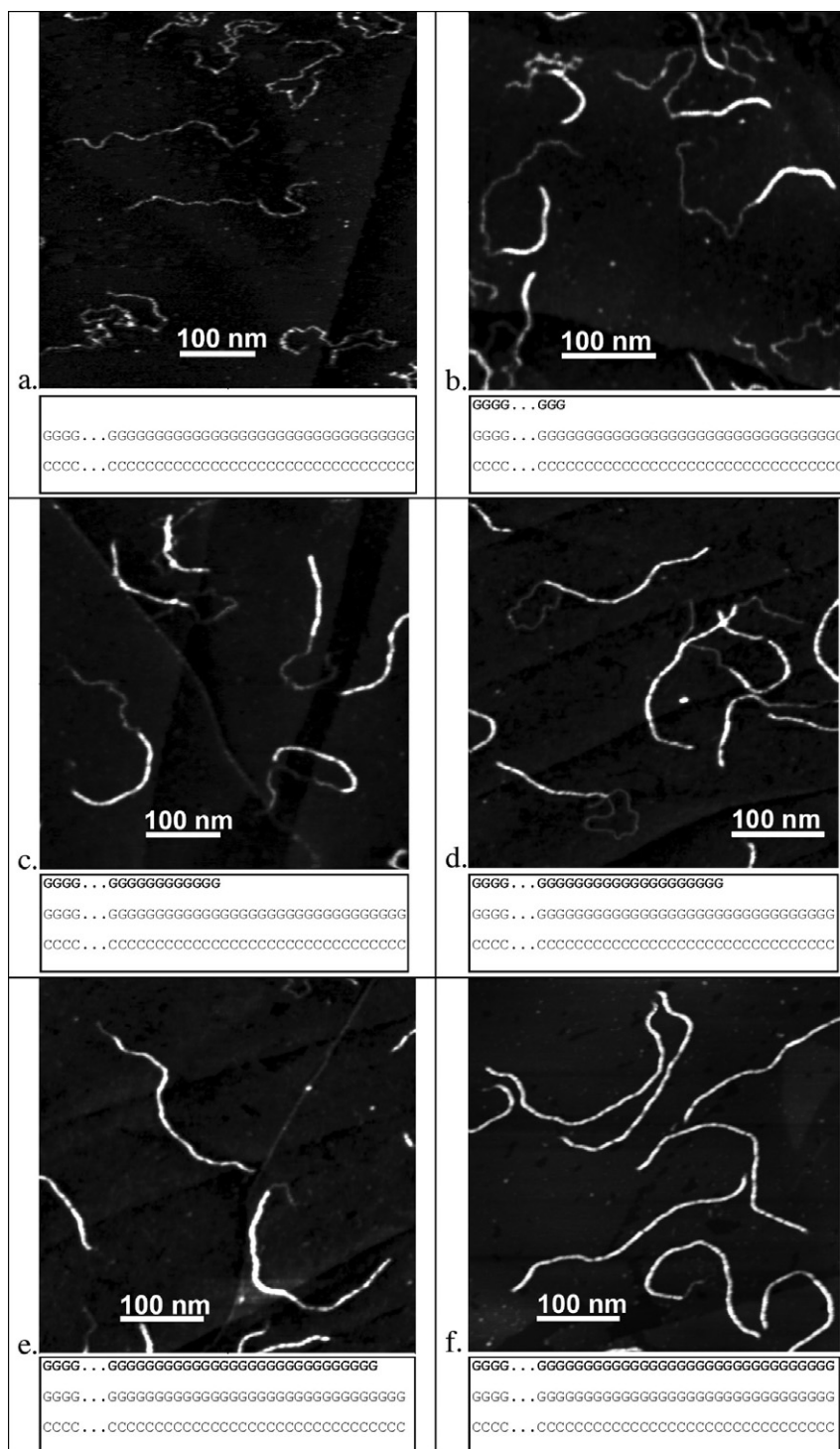


Figure 4. High-resolution 'snapshot' AFM images of the products of poly(dG-dG)-poly(dC) synthesis. The AFM images were taken after a duration of 0 h (a), 0.5 h (b), 1 h (c), 2 h (d), 3 h (e) and 6 h (f) from the beginning of the synthesis. A schematic presentation of the intermediate products of the synthesis is shown at the bottom of each picture (the G-strand *de novo* is marked in bold).

In our previous work [16] we have shown that the triplex synthesis proceeds in accordance with the slippage mechanism and includes the following steps: 1, binding of the enzyme to the 3'-end of the dG strand composing poly(dG)-poly(dC); 2, loop formation at the 3'-end of the G-strand as a result of G-base(s) addition to the strand; 3, migration of the loop through the DNA to the 5' end of the poly(dG) strand; 4, relaxation

of the loop into an overhang at the 5'-end of the poly(dG)-poly(dG)-poly(dC) motif at the distal DNA end (see figure 6). A number of successive extension and folding cycles bring the 5' end of the G-strand to a point where the synthesis has been initiated. The synthesis is slowed down at this point and further elongation of the strand does not take place in most

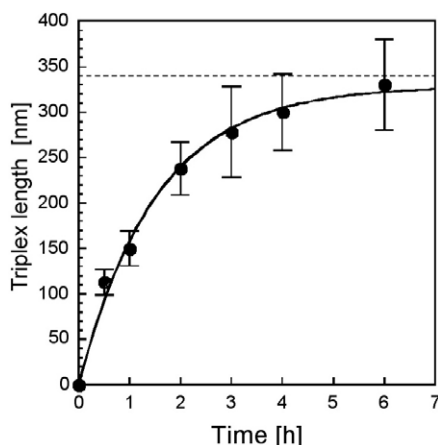


Figure 5. The dependence of the average length of the triple-stranded fragment in the synthesized molecules on the duration of synthesis. The synthesis and AFM imaging were as in figure 2. The lengths of the triple-stranded fragments were averaged for 100 molecules at each time point (circles). The error bars represent one standard deviation of the length distribution. The dotted line represents the average length of the parent poly(dG)–poly(dC) molecules.

molecules (an exception can be seen as molecule 3 in figure 2, which shows a ssDNA ‘tail’). It is clearly seen in the AFM images presented in figure 4 that the polymerase synthesis results in sequential elongation of the poly(dG) strand of the poly(dG)–poly(dC) molecule and formation of the poly(dG)–poly(dG)–poly(dC) fragment at its end. As a result the length of the triplex fragment of the DNA increases together with a reduction in the length of the double-stranded fragment as the reaction progresses, as shown in figure 5. The expansion of the strand is halted when the complete triplex is formed and both ends of the G-strands come into close proximity. Other possible synthesis mechanisms involving the simultaneous formation of triplex domains along the dsDNA or synthesis starting at both ends of the molecule can be excluded through the use of the AFM imaging technique.

A small percentage (<5%) of poly(dG)–poly(dC) was found which did not form complete triplex regardless of the

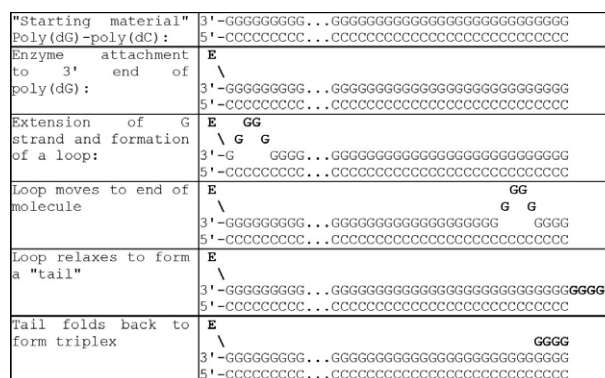


Figure 6. Schematic diagram of the synthesis mechanism of the triplex molecule. E corresponds to the enzyme, Klenow fragment of DNA polymerase I. The bold Gs represent the added nucleotides.

reaction time (see molecule 2 in figure 2). The synthesis of the triplex fragment in these molecules was apparently halted long before the complete triplex was formed. The reason for this might be the presence of mismatches in the poly(dG)–poly(dC) molecule resulting from the incorporation of an alternative base in either the poly(dG) or the poly(dC) strands.

3.3. Internal features of triplex DNA

We have also investigated the detailed molecular morphology of the triplex molecules along their axis using the same high-resolution AFM technique. Periodic variations in the height along the molecule were observed, especially in AFM phase images depicting the change of phase of the cantilever oscillations relative to its driving force (figure 7). The length of the periodic motif (an example of which is indicated in the figure by arrows) is approximately 3.4 ± 0.9 nm. This corresponds, within experimental error, to the length of periodic motif of the triplex DNA (3.9 ± 0.1 nm) determined using NMR and x-ray diffraction studies [20] as well as in contact-mode AFM studies of highly packed (liquid-crystal-like) dsDNA in a liquid cell (3.4 ± 0.4 nm) [21]. The relatively

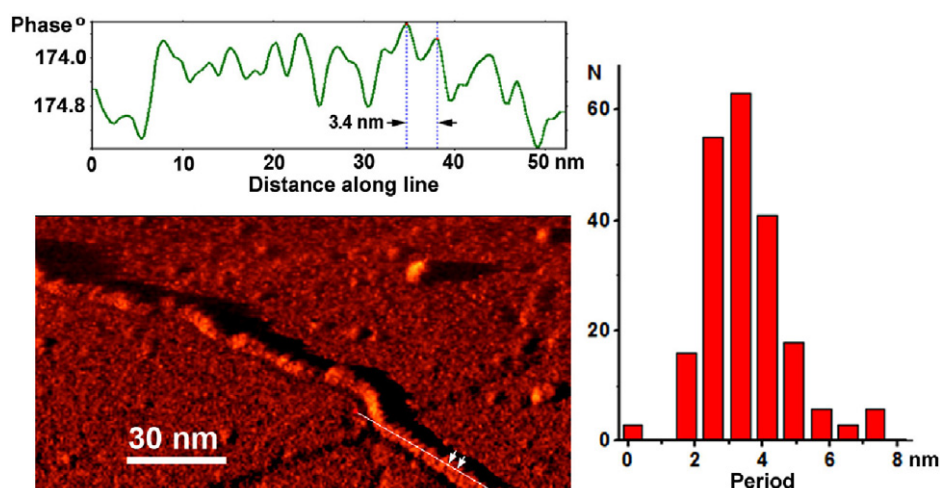


Figure 7. Left: high-resolution AFM *phase* image of poly(dG)–poly(dG)–poly(dC) triplex DNA. The top graph shows a cross-section of the image along the white line (at the bottom part of the image). Right: histogram of distances between adjacent peaks on cross-sections taken on many molecules (overall >200 values); the average distance is 3.4 nm (SD = 0.9 nm).

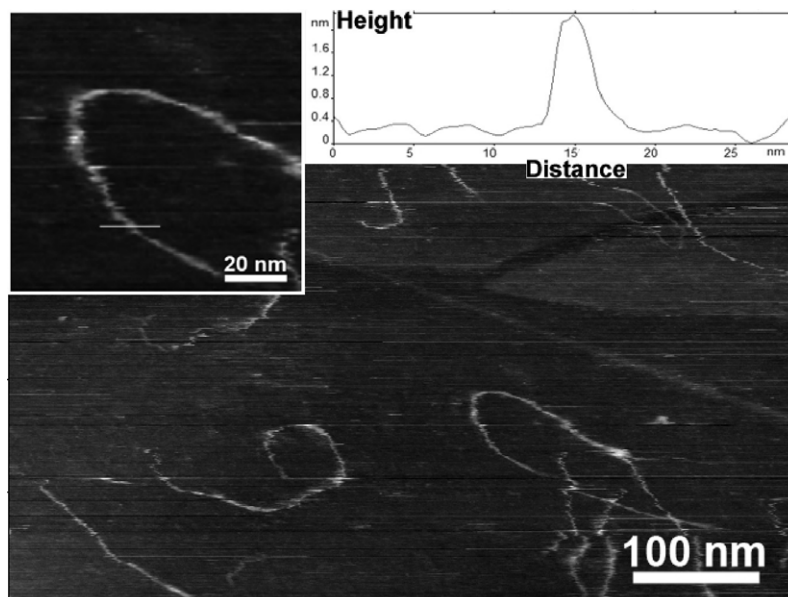


Figure 8. AFM images of poly(dG)–poly(dG)–poly(dC) triplex DNA deposited on modified HOPG and imaged under water with a high-resolution tip in tapping mode. The average width of the DNA molecule is 2.5–3 nm and its height is 2.0 ± 0.2 nm. Left inset: magnified view of a triplex molecule. Right inset: height profile of the molecule, taken along the line shown in the left inset.

wide distribution of the length of the periodic motif observed is probably due to the interaction between the molecule and the substrate, as well as to the dry ambient in which the molecule is imaged. It is worth noting that the measurements presented here are performed on single molecules and at ambient conditions. Rather than being used as a tool for measuring DNA periodicity (it is better measured by x-ray diffraction); this image illustrates the fine resolution capabilities of our imaging method.

3.4. High-resolution imaging of triplex DNA molecules under water

One of the most interesting advantages of AFM is its ability to provide high-resolution images of biomolecules under physiological conditions. We imaged the triplex DNA under water ambient (Tris buffer, pH = 7.0) in a wet AFM cell (Ntegra AFM) using a high-resolution AFM tip (figure 8). In spite of the higher noise in this image (including some horizontal fluctuations), the DNA molecule is well imaged and the average measured width of the molecule is 2.5–3 nm, as shown in the typical cross-section in figure 8. These values represent a marked improvement over standard AFM measurements made under liquid, typically showing a width of 7–10 nm [22]. Its height is 2.0 ± 0.2 nm, close to its native diameter in solution. This shows that the modified HOPG surface is hydrophilic enough to be compatible with water-based liquid cell AFM imaging, while preserving the full height of the molecules as in solution. Improved ambient conditions (e.g. acoustic screening) could probably further reduce the fluctuations in such images.

3.5. Structure of poly(dG)–poly(dC)

The synthesis of the poly(dG)–poly(dC) dsDNA itself is performed in a similar way to the synthesis of the triplex

molecule [15]. A high-resolution image of poly(dG)–poly(dC) dsDNA is shown in figure 9. The dsDNA molecules contain fragments ('loops') where the two strands are decoupled and can be seen as two single strands not associated with one another. The measured height of the decoupled strands in the loop is approximately equal to 0.4 ± 0.1 nm. Some of the poly(dG)–poly(dC) molecules contain single-strand fragments ('tails'), the height of which is also approximately equal to 0.4 nm; the height of dsDNA is ~ 1 nm, as in the previous samples. The presence of single-strand tails and loops might reflect the mechanism of poly(dG)–poly(dC) synthesis [15], which also proceeds by similar 'slippage' and includes formation of loops at the 3'-ends of each of the strands composing the poly(dG)–poly(dC) and their further migration over long molecular distances to the 5'-ends. In accordance with this mechanism, termination of the polymerase synthesis should result in relaxation of the moving loops into single-stranded fragments either at an end ('tail') or in the middle ('loop') of poly(dG)–poly(dC). These images show the advantage of the high-resolution tips, without which such high spatial resolution would not have been possible. Another interesting feature visible in figure 9 is the inhomogeneous height of double-stranded parts of the molecule. The origin of these is probably twists and kinks in the molecule, which are visible due to the high spatial resolution of the AFM tips.

4. Conclusions

The visualization of DNA by AFM with high-resolution (radius of about 1 nm) tips on modified HOPG substrates, described herein, makes it possible to detect various structural motifs along double- and triple-stranded DNA molecules as well as to resolve the fine structure of the triplex molecules. Further exploration of the molecular imaging technique described here might enable one to resolve some of the

