

Single-Molecule Reconstruction of Oligonucleotide Secondary Structure by Atomic Force Microscopy

Alice Pyne, Ruth Thompson, Carl Leung, Debdulal Roy, and Bart W. Hoogenboom*

Biomolecular structure determination generally relies on ensemble-averaging. As such, it has poor sensitivity to deviations from the average molecular structures, and usually requires crystalline samples—which for many molecules and complexes are not available. Combining soft-touch atomic force microscopy (AFM) and image analysis, we here describe a method to reconstruct the secondary structure of single extended biomolecules, without the need for crystallization. We have tested the method by accurately reproducing the dimensions of the B-DNA crystal structure. Importantly, we also resolve intramolecular variations in groove depth of the DNA double helix, which would be inaccessible for methods that rely on ensemble-averaging.

Local deformations of the DNA double helix structure are important in mediating protein–DNA binding specificity and thus in regulating gene expression.^[1] They have so far been hard to observe experimentally, in particular if the secondary structure depends on DNA supercoiling^[1b] or on local protein binding.^[1c] DNA structure is usually determined by X-ray diffraction, demonstrating, for example, few-Ångström sequence-dependent variations in depth and width of its major and minor grooves.^[2] However, X-ray diffraction is limited to providing ensemble-averaged structures based on crystals of short, linear DNA fragments. NMR,^[3] FRET^[4] and superresolution optical microscopy^[5] are complementary methods to obtain information of oligonucleotide structure in more physiological environments, but either fail to achieve the spatial resolution that is required to determine secondary structure, or are limited to the study of distances between specifically labelled positions within the molecules.

AFM occupies a unique position among such methods, because it combines the spatial resolution to resolve sub-molecular structure^[6] with the ability to image single oligo-

nucleotides of arbitrary size, complexity and conformational flexibility in physiological buffers, albeit adsorbed on a flat substrate.^[7] In some cases, it yielded sufficient resolution to resolve the helical repeat of B-DNA,^[7b,8] and recent advances in AFM technology have resulted in the visualization of both strands of the DNA double helix.^[9] To reproducibly visualize oligonucleotide secondary structure, we here minimize the invasiveness of the AFM measurements using both rapid force-distance (PeakForce Tapping)^[6b] and amplitude-modulation (tapping)^[7b] imaging modes (see the Experimental Section), and apply novel image analysis procedures to extract structural parameters.

DNA helical structure can readily be observed in our AFM images as a double-banded corrugation along the molecule (**Figure 1a**), here demonstrated for a supercoiled plasmid of about 1.2 μm length (3486 base pairs) adsorbed on mica in the presence of Ni^{2+} ions (see Experimental Section and Supporting Information Figure 1). We find that both the measured corrugation and the overall height of the plasmids critically depend on the force that the AFM tip applies to the sample. This is quantified by force-distance imaging, where continuous and rapid acquisition of force-distance curves facilitates an estimation of the peak forces during the imaging process (Figure 1b–f). At minimum peak forces, the measured height of the DNA (1.9 ± 0.2 nm) agrees with the diameter of the molecule as estimated from the B-DNA crystal structure (2.0 ± 0.1 nm; RCSB protein data bank, 1BNA).^[10] On increasing the peak force to ≈ 50 pN, the corrugation becomes marginally clearer, at the expense of few-Ångström vertical compression of the plasmid. At about 70 pN and 20% compression of the molecule, the corrugation is most visible. Beyond 100 pN, the major and minor grooves become less clear and the measured heights reduce to <1.5 nm, similar to most earlier AFM experiments in liquid.^[7b,8b] We typically observe damage to or dislocation of the plasmids on imaging at peak forces exceeding ≈ 200 pN. For peak forces up to ≈ 200 pN, the plasmid can be repeatedly compressed and then imaged again at its uncompressed height (Supporting Information Figure 2). The compression of the plasmid is also apparent from its narrower appearance (Figure 1b–d) and dislocation of a plasmid loop (see white arrow in Figure 1d) at higher forces. These results demonstrate that the oligonucleotide diameter can be accurately measured as a molecular height by AFM for low tip-sample forces.^[9]

For a comparison of the AFM images with the B-DNA crystal structure, we digitally straightened the measured surface topography of a DNA segment, using a coordinate

A. Pyne, R. Thompson, Dr. C. Leung,
Dr. B. W. Hoogenboom
London Centre for Nanotechnology
University College London
17–19 Gordon Street, London, WC1H 0AH, United Kingdom
E-mail: b.hoogenboom@ucl.ac.uk



Dr. D. Roy
National Physical Laboratory
Hampton Road, Teddington, TW11 0LW, United Kingdom

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/sml.201400265

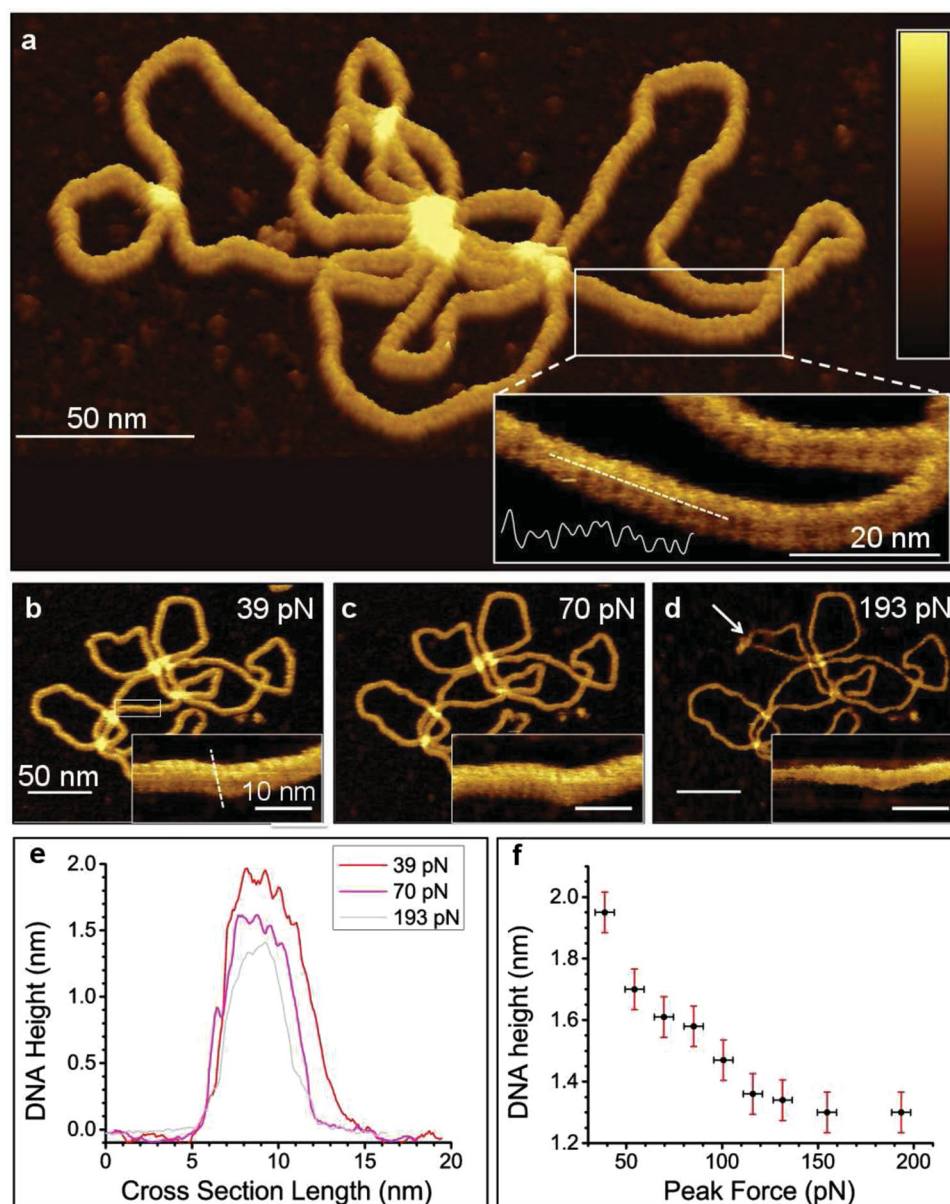


Figure 1. Double-helix corrugation and height of a DNA plasmid in AFM topography. (a) A plasmid imaged at a peak force of 49 pN. Inset: High-resolution AFM topography of the DNA, and a height profile taken with a 5 pixel (≈ 0.5 nm) width along the dashed line, showing major and minor grooves. Color scales: 3.5 nm; 1.4 nm (inset). (b–d) A plasmid imaged at peak load forces of 39, 70, and 193 pN, respectively, with the major and minor grooves of the DNA double helix visualized at higher magnification (insets). Color scales: 3 nm (for low magnification); 2 nm (for the insets). (e) Height profiles measured across the DNA, as marked on the inset of b by a dashed line, for different peak forces. (f) Measured height along the same section across the molecule (as e), as a function of peak force.

transformation to a system in which one coordinate (s) represents the position along the longitudinal axis of the molecule, and the other coordinate (q) the distance from this central axis (see Experimental Section and Supporting Information Figure 3), i.e., $q = 0$ on the longitudinal axis of the molecule. The periodic double-banded structure—much clearer at this higher magnification and enhanced color contrast (Figure 2a–b)—corresponds to the two oligonucleotide strands of the double helix, and the tilt of the banding marks the right-handedness of the B-DNA secondary structure. The helical periodicity can also be identified from the power spectra (Figure 2c) obtained by a one-dimensional Fourier transform along the horizontal axis in Figure 2b. Two pronounced peaks

are observed, corresponding to the first ($2\pi/\lambda$) and second ($4\pi/\lambda$) order wave vectors, which each yield a helical repeat of $\lambda = 3.6 \pm 0.2$ nm, in excellent agreement with the crystal structure. The smallest discernible peak in the experimental power spectrum corresponds to the third order wave vector ($6\pi/\lambda$), indicating a spatial resolution of 1.2 ± 0.2 nm. This lateral resolution is similar to the best lateral resolution that has been obtained by AFM on arrays of membrane proteins,^[6c] which due to their flatness and two-dimensional periodicity are much more amenable to high-resolution imaging.

Remarkably, the double-helix contrast shows significant variations along the molecule (Figure 2a–b). These variations are reproducible between trace and retrace scans as well as

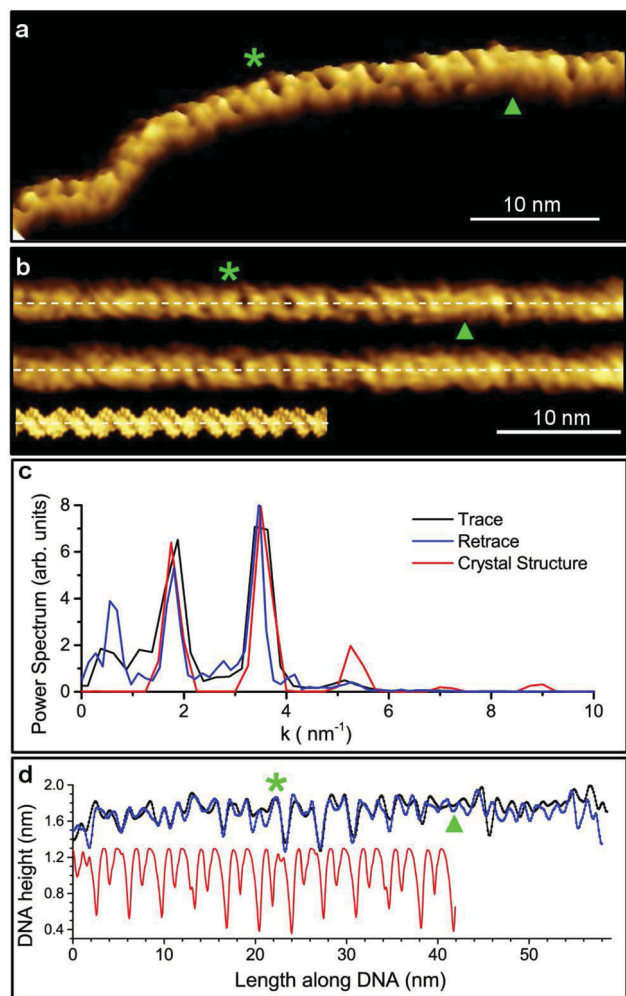


Figure 2. Helical repeat, right-handedness, and groove depths of B-DNA. Green asterisks mark a deep major groove, and green triangles a shallow one. (a) AFM topography of plasmid DNA, displayed at a (saturated) color scale of 1.1 nm. (b) Digitally straightened trace (top) and retrace (bottom) image of a, compared with a space-filling representation of the B-DNA crystal structure. (c) Power spectra for the images in b, obtained by a Fourier transform in the horizontal direction. The black and blue lines refer to the trace and retrace AFM images, respectively; the red trace to a similarly processed image of the crystal structure. (d) Height profiles taken along the dashed lines in b averaged over a 5 pixel (≈ 0.5 nm) width. The profile of the crystal structure has been offset by -0.8 nm for clarity.

between subsequent images, are independent of the image analysis procedure, and can be quantified as variations of the major groove depth in height profiles measured along the molecule (Figure 2d and Supporting Information Figure 4). The major and minor grooves of the double helix appear as, respectively, 0.1–0.7 nm and ≈ 0.1 nm depressions, which on average are reduced compared with their 0.8 nm and 0.4 nm depths in the crystal structure^[10] (as estimated from a profile measured along its space-filling representation, Figure 2d, red line). This suggests that the AFM tip is too large to fully penetrate the grooves, implying a limitation for the accuracy at which AFM can measure the absolute groove depths. We interpret the measured (relative) variations in the major groove depths as structural variability possibly induced by

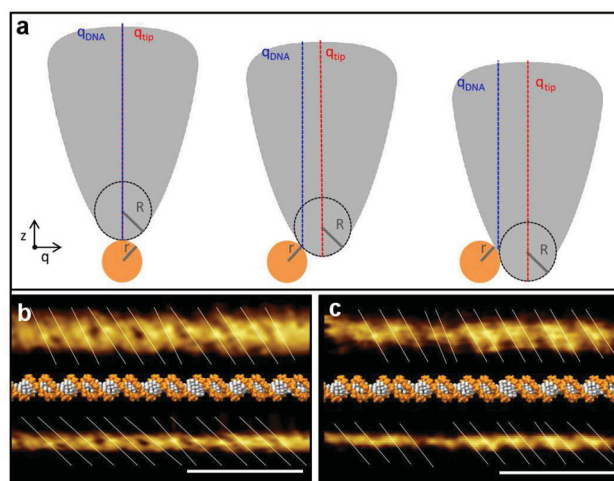


Figure 3. Measuring the chiral angle of the double helix. (a) Schematic showing the AFM tip of radius R at different positions with respect to a cylindrical molecule of radius r (viewed as cross section). (b, c) AFM topography of digitally straightened DNA segments before (top) and after (bottom) correction for the finite tip size, compared with the B-DNA crystal structure, in two different experiments. $r = 1.1 \pm 0.1$ nm, $R = 1.3 \pm 0.2$ nm (b); and $r = 0.9 \pm 0.1$ nm, $R = 1.7 \pm 0.2$ nm (c). Scale bars: 10 nm; color scales: 1.1 nm. Straight lines in the AFM images mark the strands of the double helix that were used to determine the average chiral angle with respect to the longitudinal axis of the molecule.

the torsional stress in the supercoiled plasmid,^[1b] and stabilized by the Ni^{2+} -mediated interaction with the mica substrate. Importantly, these variations demonstrate the power of our method in not only accurately reproducing average secondary-structure dimensions, but also their variability within a single molecule.

The AFM images thus provide an accurate representation of periodic and non-periodic features along the molecule, and show groove depths that approach the idealized structure (as would be measured on the B-DNA crystal structure with an infinitely sharp and sensitive AFM probe) to within a few Ångströms. However, the width of the DNA is largely overestimated due to the convolution with the finite-size AFM tip (Figure 1e and Figure 3a). As a result, the chiral angles of oligonucleotide secondary structure (as measured with respect to the longitudinal axis of the molecule) are overestimated as well (Figure 3b-c).^[9b]

To obtain accurate values of the chiral angles of the DNA double helix, we first determine the molecular radius (r) and tip radius (R) from the measured AFM topography, and next use these to correct the AFM images for the finite size of the tip (see Experimental Section, and Ref. [11] for a similar analysis of scanning tunnelling microscopy data on carbon nanotubes). For the higher-quality images, tip radii are typically found to be 1–2 nm (below nominal manufacturers' specifications, see Supporting Information Table 1). Figure 3b-c show examples of the correction procedure for data acquired in two different experiments. Average chiral angles (\pm standard deviation) are corrected, respectively, from $23 \pm 4^\circ$ to $36 \pm 4^\circ$ and from $21 \pm 3^\circ$ to $37 \pm 5^\circ$ for Figure b and c, in excellent agreement with the $36 \pm 1^\circ$ predicted from the B-DNA crystal structure.

The comparison with the B-DNA crystal structure demonstrates that our method provides an accurate reconstruction

of secondary structure based on a single oligonucleotide molecule (see Supporting Information Table 2 for a summary of structural parameters). Via the measured groove depths, it detects intramolecular variability that is inaccessible to other structural methods as they usually rely on ensemble-averaging—with potential repercussions for protein and ligand binding during gene expression and replication.^[1a] Based on commercially available equipment, our method is sufficiently general and accessible to have wider applicability, e.g., for the determination of yet unsolved structures of coiled-coil protein complexes^[12] and for the structural verification of engineered oligonucleotide structures such as DNA origami.^[13] Finally, given the here quantified interdependence of load force, tip size, and measured AFM topography of known biomolecular structure, our method also lends itself exquisitely for in-situ benchmarking of AFM systems and probes.

Experimental Section

Sample Preparation: Double-stranded plasmid DNA was physisorbed on mica in the presence of Ni^{2+} ions to bridge the negative surface charges of the mica substrate and the DNA molecules in solution.^[7b] Once adsorbed, the DNA appeared in a two-dimensional projection of three-dimensional superstructure, showing various degrees of supercoiling. Freshly cleaved mica (Agar Scientific, UK) was covered with 25 μL of 20 mM NiCl_2 (Sigma-Aldrich, UK) solution, followed by addition of 25 μL of 10 mM HEPES pH 7.0 (Sigma-Aldrich) buffer, and of 7 μL of 3 $\mu\text{g}/\text{mL}$ 3486 base-pair plasmid DNA solution (pmaxGFP nucleofector kit, Lonza, Basel, Switzerland; characterized by agarose gel electrophoresis as described elsewhere).^[9a] All solutions were allowed to equilibrate at room temperature before the sample preparation, to reduce subsequent drift in the microscope. After 30 minutes, the solution was diluted by addition of 50 μL of 10 mM HEPES pH 7.3. This reduced the NiCl_2 concentration to 5 mM NiCl_2 , which implies a Debye screening length of ≈ 2.5 nm. All AFM measurements were performed in liquid under these conditions. At lower NiCl_2 concentrations, the DNA was more loosely bound to the surface, compromising spatial resolution at the imaging speeds used in these experiments. At higher concentrations, NiCl_2 showed stronger precipitation on the surface, which compromised tip stability and prevented accurate height measurements of the DNA (Supporting Information Figure 1).

AFM imaging: Rapid force-distance (PeakForce Tapping) imaging was carried out on Multimode 8 and FastScan Bio AFM systems (Bruker, Santa Barbara, CA, USA).^[6b] In these experiments, continuous force-distance curves were recorded and the tip-sample feedback was set by the peak force as referenced to the force baseline. The following cantilevers were used: MSNL-F (Bruker), biolever mini (Olympus, Tokyo, Japan), and FastScan DX (Bruker), see Supporting Information Table 1. Force-distance curves were recorded over 10 nm (PeakForce Tapping amplitude of 5 nm), at a frequency of 4 kHz. The PeakForce setpoint for approach was set to 0.1 V, corresponding to a deflection of a few nm and a force of the order of 0.1 nN. The deflection sensitivity was calibrated from a force curve with the same (relative) setpoint, after which the spring constant was calibrated from the thermal noise of the cantilever away from the sample surface. Measured

deflections could thus be converted to forces. For imaging, the vertical scan limit was reduced to 1 μm .

Amplitude-modulation (AC mode/tapping)^[7b] experiments were carried out on a Cypher microscope (Asylum Research, Santa Barbara, CA, USA) with biolever-mini cantilevers (Olympus). The cantilevers were piezo-acoustically driven at an amplitude of 1.3 nm at the resonance frequency of the cantilever. This resonance frequency (≈ 27 kHz) was determined from the thermal noise spectrum of the cantilever. The amplitude setpoint for approach was set at about 70% of the free amplitude. At these settings, several false approaches were detected due to variations in the effectiveness of the piezo-acoustic actuation during the coarse tip-sample approach. Drive amplitude and frequency were adjusted at this point, until tip-sample contact was established. For imaging, the amplitude setpoint was increased to just below the value at which the cantilever was lifted off the sample.

After approach of the AFM tip to the sample, the scan area was set to $1 \times 1 \mu\text{m}^2$ to locate plasmids that had adsorbed to the mica without excessive supercoiling, such that roughly straight segments were exposed over several tens of nanometres. Higher-resolution images were taken at scan sizes of $100 \times 100 \text{ nm}^2$ or less. Images were recorded at 512×512 pixels at a line rate of 2.5 Hz (Bruker Multimode 8 and Fast Scan Bio) or 1.5 Hz (Cypher). Each scan line in the image was scanned from left to right (trace) and from right to left (retrace). The observed topographic features were verified for their consistency between trace and retrace images (Figure 2 and Supporting Information Figure 4), as well as for their reproducibility in subsequent scan frames (Supporting Information Figure 4).

Similar AFM topography images were obtained in PeakForce Tapping on the Bruker instruments and in amplitude modulation on the Cypher. For quantifying the peak force (Figure 1), however, we preferred PeakForce Tapping, as the quantification of imaging forces in amplitude modulation was less straightforward. The data in Figure 1 were recorded on the Bruker Multimode 8 using PeakForce Tapping, the data in Figure 2 and 3b on the Bruker Fast Scan Bio using PeakForce Tapping, and the data in Figure 3c on the Asylum Cypher AFM system with amplitude modulation.

Image analysis: A dedicated Matlab code (MathWorks, Natick, Massachusetts) was written to load the AFM images as ASCII files, trace the DNA, and carry out a transformation from the xyz coordinates of the surface topography to sqz coordinates, where the unit vector \hat{s} refers to the direction along the contour of the DNA, and \hat{q} to the direction perpendicular to \hat{s} at all points (see Supporting Information Figure 3 for a pictorial representation). For any particular position s along the DNA, the q coordinate of a nearby pixel was determined from the vector product $\vec{p} \times \hat{s}$ where \vec{p} was the vector from the position defined by s to the position of the nearby pixel. The surface height z was not affected by this transformation.

The DNA was traced, starting from a user-defined point on the DNA, by defining (q, z) points with respect to a segment of set length s typically about 2 nm along a trial direction \hat{s} ; fitting of these nearby (q, z) points with a Gaussian superposed to a linear background; defining the peak position of this Gaussian as a new point on the DNA contour; defining a new trial direction from the line connecting the previous two points; and repeating this procedure until a user-defined end-point was reached. The DNA contour was then traced by a smooth spline along thus defined positions.

In the next step, this DNA trace was divided into s values at a resolution corresponding to the pixel size in the original image, and all (x, y) points in the images (Supporting Information Figure 3) attributed to the nearest s position on the DNA contour. Finally, the resulting (s, q, z) coordinates were interpolated to yield an image with equally spaced (s, q) pixels.

Line-by-line background subtraction of raw AFM images, cross-sectional and profile analysis, as well as color contrast adjustments, Gaussian filtering with a full-width half maximum of 2 pixels, and Fourier analysis were carried out by the open-source image analysis program Gwyddion.^[14] Graphical representations of the B-DNA crystal structure were generated using the UCSF Chimera package.^[15]

Correction for finite tip size: Assuming a cylindrical molecule and a spherical AFM tip (with radii defined as r and R , respectively; see Figure 3a), the lateral position of the point of contact between a sample and tip (q_{DNA}) relates to the lateral position of the tip (q_{tip} , i.e., the measured position in an AFM experiment) via

$$q_{DNA} = \frac{q_{tip}}{1 + R/r} \quad (1)$$

which follows from simple geometrical arguments. In our experiments, r was estimated from accurate, low-force measurements of the maximum height of the molecule ($=2r$), and R from its measured full-width half maximum ($\equiv 2q_{tip,1/2}$):

$$R = \frac{(q_{tip,1/2})^2 - r^2}{2r} \quad (2)$$

Here, $q_{tip,1/2}$ was determined from a fit with a Gaussian peak function, as more complicated and accurate descriptions of the cross-sectional profile did not lead to significantly different results. Using Equations (1), (2), the straightened oligonucleotide AFM images in sqz coordinates were thus corrected for the effects of tip convolution.

Coordinates with $q^*_{tip} > 2\sqrt{rR}$ were excluded from the analysis, as these correspond to tip positions where the tip is expected to be in contact with the substrate, and no longer with the DNA. Consequently, the DNA surface is only probed over a width $2q^*_{DNA} = 4\sqrt{rR} / (1 + R/r) \leq 2r$. In other words, though tips with larger sizes yield a wider appearance of the plasmid, they will probe a narrower part of the molecular surface. Using similar geometrical arguments, one may calculate a corrected height $z_{DNA} \geq z_{tip}$, where z_{tip} refers to the very end of the tip. However, unlike the correction of the lateral dimension q , this height correction in our experiments was not significant compared to the noise, and therefore was not pursued any further.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

The authors would like to thank Bruker (Santa Barbara, CA, USA) for technical support and for providing instrumentation, particularly Andrea Slade, Bede Pittinger, Steve Minne, Shuiqing Hu and Chanmin Su for their help and informative discussions; and Stefan Howorka for proofreading the manuscript. This work was funded by the UK Biotechnology and Biological Sciences Research Council (BB/G011729/1) and Engineering and Physical Sciences Research Council (EP/G036675/1). Alice Pyne and Debdulal Roy acknowledge financial support from the ChemBio Programme of the National Measurement System, UK. These Acknowledgements were updated on August 27, 2014.

- [1] a) R. Rohs, X. Jin, S. M. West, R. Joshi, B. Honig, R. S. Mann, *Annu. Rev. Biochem.* **2010**, *79*, 233–269; b) J. M. Fogg, G. L. Randall, B. M. Pettitt, D. W. L. Summers, S. A. Harris, L. Zechiedrich, *Q. Rev. Biophys.* **2012**, *45*, 257–299; c) S. Kim, E. Broströmer, D. Xing, J. Jin, S. Chong, H. Ge, S. Wang, C. Gu, L. Yang, Y. Q. Gao, X.-d. Su, Y. Sun, X. S. Xie, *Science* **2013**, *339*, 816–819.
- [2] C. Oguey, N. Foloppe, B. Hartmann, *PLoS One* **2010**, *5*, e15931.
- [3] E. N. Nikolova, E. Kim, A. A. Wise, P. J. O'Brien, I. Andricioaei, H. M. Al-Hashimi, *Nature* **2011**, *470*, 498–502.
- [4] T. Fessl, F. Adamec, T. Polivka, S. Foldynova-Trantirkova, F. Vacha, L. Trantirek, *Nucleic Acids Res.* **2012**, *40*, e121–e121.
- [5] C. Flors, W. C. Earnshaw, *Curr. Opin. Chem. Biol.* **2011**, *15*, 838–844.
- [6] a) H. Asakawa, K. Ikegami, M. Setou, N. Watanabe, M. Tsukada, T. Fukuma, *Biophys. J.* **2011**, *101*, 1270–1276; b) Y. F. Dufrière, D. Martínez-Martín, I. Medalsy, D. Alsteens, D. J. Müller, *Nat. Methods* **2013**, *10*, 847–854; c) A. Engel, H. E. Gaub, *Annu. Rev. Biochem.* **2008**, *77*, 127–148.
- [7] a) Y. L. Lyubchenko, L. S. Shlyakhtenko, T. Ando, *Methods* **2011**, *54*, 274–283; b) H. G. Hansma, *Annu. Rev. Phys. Chem.* **2001**, *52*, 71–92.
- [8] a) M. Maaloum, A.-F. Beker, P. Muller, *Phys. Rev. E* **2011**, *83*, 031903; b) J. Mou, D. M. Czajkowsky, Y. Zhang, Z. Shao, *FEBS Lett.* **2009**, *371*, 279–282; c) S. Santos, V. Barcons, H. K. Christenson, D. J. Billingsley, W. A. Bonass, J. Font, N. H. Thomson, *Appl. Phys. Lett.* **2013**, *103*, 063702.
- [9] a) C. Leung, A. Bestembayeva, R. Thorogate, J. Stinson, A. Pyne, C. Marcovich, J. L. Yang, U. Drechsler, M. Despont, T. Jankowski, M. Tschöpe, B. W. Hoogenboom, *Nano Lett.* **2012**, *12*, 3846–3850; b) S. Ido, K. Kimura, N. Oyabu, K. Kobayashi, M. Tsukada, K. Matsushige, H. Yamada, *ACS Nano* **2013**, *7*, 1817–1822.
- [10] H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson, *Proc. Nat. Acad. Sci. USA* **1981**, *78*, 2179–2183.
- [11] L. Venema, V. Meunier, P. Lambin, C. Dekker, *Phys. Rev. B* **2000**, *61*, 2991–2996.
- [12] D. N. Woolfson, G. J. Bartlett, M. Bruning, A. R. Thomson, *Curr. Opin. Struct. Biology* **2012**, *22*, 432–441.
- [13] M. Erkelenz, D. M. Bauer, R. Meyer, C. Gatsogiannis, *Small* **2013**, *10*, 73–77.
- [14] D. Nečas, P. Klapetek, *Cent. Eur. J. Phys.* **2012**, *10*, 181–188.
- [15] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.

Received: January 29, 2014
 Revised: March 20, 2014
 Published online: April 17, 2014