Nanoscale structural features determined by AFM for single virus particles

Shu-Wen Chen, Michael Odorico, Matthieu Meillan, Luc Vellutini, Jean-Marie Teulon, Pierre Parot, Bernard Bennetau, Jean-Luc Pellequer

Nanoscale

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Nanoscale structural features determined by AFM for single virus particles

Cite this: *Nanoscale*, 2013, **5**, 10877

Shu-wen W. Chen,[ab] Michael Odorico,[a] Matthieu Meillan,[c] Luc Vellutini,[c] Jean-Marie Teulon,[a] Pierre Parot,[a] Bernard Bennetau[c] and Jean-Luc Pellequer*[a]

In this work, we propose "single-image analysis", as opposed to multi-image averaging, for extracting valuable information from AFM images of single bio-particles. This approach allows us to study molecular systems imaged by AFM under general circumstances without restrictions on their structural forms. As feature exhibition is a resolution correlation, we have performed AFM imaging on surfaces of tobacco mosaic virus (TMV) to demonstrate variations of structural patterns with probing resolution. Two AFM images were acquired with the same tip at different probing resolutions in terms of pixel width, *i.e.*, 1.95 and 0.49 nm per pixel. For assessment, we have constructed an *in silico* topograph based on the three-dimensional crystal structure of TMV as a reference. The prominent artifacts observed in the AFM-determined shape of TMV were attributed to tip convolutions. The width of TMV rod was systematically overestimated by ~10 nm at both probing resolutions of AFM. Nevertheless, the effects of tip convolution were less severe in vertical orientation so that the estimated height of TMV by AFM imaging was in close agreement with the *in silico* X-ray topograph. Using dedicated image processing algorithms, we found that at low resolution (*i.e.*, 1.95 nm per pixel), the extracted surface features of TMV can be interpreted as a partial or full helical repeat (three complete turns with ~7.0 nm in length), while individual protein subunits (~2.5 nm) were perceivable only at high resolution. The present study shows that the scales of revealed structural features in AFM images are subject to both probing resolution and processing algorithms for image analysis.

Received 24th May 2013
Accepted 12th August 2013

DOI: 10.1039/c3nr02706f

www.rsc.org/nanoscale

1 Introduction

Atomic force microscopy (AFM) is now extensively used for imaging sample surfaces in a scale range of micrometer (μm) to nanometer (nm). Despite the increasing growth of its use, AFM was employed mainly in solid-state physics, particularly in materials research, while to a less extent in biology.[1–3] Equipped with improved instrumentation setups and image treatments, AFM has evolved into an imaging tool which may probe biological specimens such as proteins at a resolution of sub-nanometer.[4] In addition, AFM is applicable to force measurements of inter-molecular interactions that play a critical role in understanding biological activities.[5,6] Moreover, this technique can be combined with other optical imaging techniques for identification of molecular components in a biological system. Lastly, AFM has no stringent constraints on sample preparations like X-ray crystallography or two-dimensional electron microscopy (2D EM); it enables us to study surface structures of biological particles under physiological conditions.[7]

Recently, it has been shown that a crystal structure of biological macromolecule can be successfully docked into the corresponding AFM topographic envelope which was acquired at a lower resolution than the atomic scale.[8] This indicates that an AFM topograph can be used as a guideline for patching multiple components of a macromolecule together,[9,10] analogous to EM envelopes wrapping X-ray structures.[11] Nevertheless, an issue was raised on the AFM imaging resolution claimed in the literature. The resolving power of AFM on surface features of a macromolecule should be compared with other techniques such as X-ray crystallography, NMR or EM. Naturally, a molecule with available 3D atomic coordinates becomes a favorite object for testing AFM resolving power on molecular structures. However, a direct comparison still remains challenging for at least one trivial fact that sample conditions are different among various measuring instruments. For example, one may have to freeze the sample for cryo-EM or to crystallize the molecule for X-ray crystallography. For AFM imaging, a deliberate adhesion between the sample and substrate material is indispensable for depositing the studied molecules over the probed plate.

As a bridge for exploring structural attributes of biological particles at a level between atomic and assemblage scales,

[a]*CEA, iBEB, Service de Biochimie et Toxicologie Nucléaire, DSV/iBEB/SBTN – Bat 170, BP17171, F-30207 Bagnols sur Cèze, France. E-mail: jlpellequer@cea.fr; Fax: +33 (0) 466 79 19 05; Tel: +33 (0)466 79 19 43*

[b]*13 Avenue de la Mayre, 30200 Bagnols sur Cèze, France*

[c]*Institut des Sciences Moléculaires (UMR 5255 - CNRS), Université Bordeaux 1, 351 Cours de la Libération, 33405 Talence, France*

high-resolution AFM imaging was commonly thought to distinguish a surface feature on a scale equal or better than 20 Å (~2 nm).[12] However, when observations go down to a very fine scale, thermal motions or intrinsic fluctuations of molecular structures become evident, particularly in the region of loops in a protein. It leads to conformational uncertainties encountered in many structure determination techniques. Some 3D atomic coordinates of molecules deposited in PDB are represented by an ensemble of structures to take into account such structural variations. Multiple conformations have been shown to be advantageous over single structures on predictions of disulfide-bonded sites in a molecule.[13] Nevertheless, many of the structural analyses relied on one representative structure of molecule for better visual percept and easier data manipulations. In cryo-EM, a strategy called iterative real-space helical reconstruction (IRSHR)[14] was developed for treating structural disorders and distortions of helical specimens to obtain single 3D structures. In AFM imaging, multi-image averaging[15] or a 2D-Gaussian fitting of cross-correlation peaks[9] has been used to improve the signal-to-noise ratio for topographic representations. Indeed, noise reduction did enhance individuality and visibility of imaged objects, and thus improved the apparent resolution of the AFM image.[16]

However, multi-image averaging sometimes was found blurring, instead of enhancing, the image contents even for a 2D crystal-like system[12] by the fact that the revealed structural features of molecules cannot be fully guaranteed to be well aligned with one another from the series of AFM images. This structure misalignment can be partially attributed to the conformational fluctuations mentioned earlier. These issues are even more exhorted in high-speed AFM imaging which often causes line skipping and provokes unequal acquisition quality within the same image or in time-series images.[17,18] In most occasions, high-resolution AFM imaging is performed on rigid and well-ordered samples such as 2D crystals or highly packed membrane proteins,[19,20] which are not always generalizable. Moreover, convolution effects of probing tip and applied forces are known to be critical to jeopardize characterization of the molecular shape. All of these problems lead to limited usefulness of AFM imaging and the difficulty for the image quality to meet the high-resolution standard.

In order to advance applications of AFM imaging on molecular topography, feature analysis on a single image (hereafter called single-image analysis) is an indispensable step for extracting biological information under a general condition.[21] In single-image analysis, each image is considered as an independent source of information for the biological system under study. One may synthesize all the contents in individual images for later use (*e.g.* frame averaging) or gain insights straightaway into molecular organizations from a particular image. The latter is of great interest in studying dynamic behaviors of molecules by AFM imaging where each image represents a unique configuration of molecules in time–space dimension. To determine feature patterns by single-image analysis, we attempted to understand, within previously mentioned limitations, to what extent the desirable information can be extracted from an AFM image, and how a 3D molecular structure can be related to the 2D information therein or *vice versa*. We were also concerned about how the visibility of a surface feature of a bio-particle imaged by AFM can be influenced by the probing resolution processing tools used in the single-image analysis.

To achieve the above aims, we chose the tobacco mosaic virus (TMV) as a demonstration specimen for the present study. The organization of TMV assembly has been widely studied using imaging techniques such as AFM and EM.[22,23] The 3D crystal structure of TMV was determined by X-ray fiber diffraction, revealing the TMV assembly in the form of a right-handed helix.[24,25] Further refinements on the 3D atomic model of the helical TMV assembly were performed using cryo-EM.[26] The regularity of the TMV structure simplifies the comparison of results out of single-image analysis on AFM images with that using the X-ray technique. In the present work, we have imaged TMV particles by AFM at two probing resolutions to study the impact of resolution on variations of observed surface features. On each raw image, we have analyzed the data and manifested structural patterns embedded in the measured surfaces of TMV. In the next section, we describe the details of AFM instrumental setups and computational procedures for data processing, followed by the results and discussion as well as conclusions.

2 Experimental section

In the present work, the single-image analysis on AFM topography was focused on geometric properties and structural patterns of single particles of TMV. On determination of the geometric shape and feature extraction for a bio-particle, we pursued the goal that surface features revealed in the AFM image can be resolved at a level of high resolution. To investigate the impact of probing resolution on the apparent resolution for a 2D molecular structure, the first task was to acquire AFM images at various probing resolutions. In this work, two probing resolutions were established in AFM imaging to probe TMV surfaces.

Sample preparation and AFM instrumentation

TMV particles were deposited over the synthetic polymers called self-assembled monolayers (SAMs) on the substrate plate. The characteristic of the SAM structure is that it contains a long alkyl chain of about 2.5–3.0 nm grafted on silicon wafer with one end while the carboxylic group of the other end is exposed on the sample surface.[27] The concentration of TMV was 15 mg ml^{-1} in a solution of 1 mM EDTA at pH 7.7. The surface of SAM was covered with one drop of 130 μl HEPES buffer (in 10 mM; 140 mM KCl at pH 7.4). 20 μl of TMV were then inserted into the HEPES drop and incubated for 30 min. Prior to imaging, the SAM surface was gently rinsed with pure water (Millipore) and later dried using a Laboport vacuum pump (KNF Neuberger, Trenton, NJ, USA). The sample was kept at room temperature during the imaging operation.

AFM imaging was performed in air using ScanAsyst™ (PeakForce Tapping™ mode) with a Multimode AFM (Bruker AXS, Santa Barbara, USA) operated by a NanoScope® V

controller. The only parameter controlled by ScanAsyst software was the gain; all other ScanAsyst parameters were manually controlled. The peak force ramp was set to 10 nm and all other peak force parameters were set as default. During imaging, the peak force set point was manually adjusted to obtain the best visual imaging at the lowest imaging force. A silicon tip was attached to the silicon nitride lever (SNL, Bruker AFM Probes, Santa Barbara, USA) for probing the sample surface; the spring constant of the cantilever was in a typical range, 0.12–0.35 N m^{-1}. The physical size of the image was set to 1 μm and 250 nm for two square AFM topographs, respectively. The former image was scanned at 90° whereas the latter image was scanned at 37° to obtain a horizontally oriented virus particle. Thus, both images represent the same virus particle that has been scanned by the same tip within 60 min of time. The scan frequencies were adjusted to 1 and 2 Hz per scan line with respect to the image sizes of 1 μm and 250 nm. The probing points per line were 512 for both images. For convenience, we refer the first image to the low resolution while the second one to the high resolution. As a result, the pixel width of high resolution (4.88 Å) is four times finer than that of low resolution.

Feature extraction from topographic images

Since diverse forms of TMV assembly have been studied and published elsewhere,[28,29] the geometric description of the virus particle aims to determine its lateral size and height. In measuring the TMV size, the width or diameter of the rod-shaped virus was concerned instead of the length of a single TMV particle, which is controlled by the length of its single-stranded RNA. For analyzing topographic contents in the low-resolution AFM image, a partial segment of TMV particle was cropped. The following describes the computational procedures for geometric characterization of the TMV particle.

Segmentation of single TMV particles from AFM raw images. Segmentation involves defining object boundaries. The area encircled by boundary pixels was called the segmented region while all others were labeled as non-segmented regions. The boundary pixels were identified using the zero-crossing detector which calculates the Laplacian of Gaussian (LOG) function at each pixel of the image.[30] In the present zero-crossing detector, a pixel was qualified as a zero-crossing point by thresholding the difference of opposite-signed LOG values of two adjacent pixels to that pixel; the thresholding parameter was denoted as Δ. In addition to Δ, two parameters, N_{mask} and σ, were implemented for tuning the spread of Gaussian smoothing function. With same values for N_{mask} and σ, the distribution of Gaussian function would be identical in space.

Estimation of TMV rod diameter. The segmented region is where TMV intensities are distributed. First, we calculated the eigenvectors and eigenvalues of the covariance matrix formed of TMV intensities, where the origin was located at the weighted centroid and the intensity was used as the weighting factor. As a result, the TMV envelope can be characterized by the two eigenvectors of which one provides us with information about the line for best fitting the data. The other eigenvector gives us a deviation account of data from the fitting line. Intuitively, the former vector indicates the longitudinal direction of the virus rod, while the latter specifies the transversal direction.

The diameter of TMV rod can be crudely obtained from width or height measurements of the TMV envelope. The width was measured as the cross length of the virus rod in parallel to the transversal direction. The height value was considered as the net value of total height subtracting the supporting height or the height of SAMs. Owing to the lack of knowledge on how much SAMs would bend under the weight of TMV, the supporting height was estimated in an arbitrary manner as the average of the bottom 1% of intensities in non-segmented regions. Oppositely, the total height was considered as the average of the top 1% of intensities in the segmented region. Alternatively, we may also derive the diameter of TMV by non-linear fitting of cross-sections of the virus rod to ellipses, where the major and minor radii of the ellipse are two fitting parameters. Actually, two times of the major radius are equal to the horizontal diameter (width-like), while two times of the minor radius correspond to the vertical diameter (height-like). When the values of two radii are identical, the fitted geometry becomes a circle.

To make surface features of TMV visible in the AFM image, the intensity processing involved reduction of stripe noises,[16] histogram equalization,[31] and calculations of mixed partial derivative of pixel intensity, *i.e.* $D_{xy}{}^2 I(x,y) \equiv \partial^2 I(x,y)/\partial x \partial y$, where $I(x,y)$ is the intensity at pixel (x,y). Histogram equalization is a means of intensity transformation by distributing pixels evenly in the intensity scale for enhancing the image contrast. We chose this tool for its simplicity and effectiveness that the implemented program (called the histogram equalizer) requires no other input parameter than the bin number of histogram. Systematically, we set the bin number to 128. The $D_{xy}{}^2$ operation was used for expressing the magnitude of intensity change.

Construction of the *in silico* X-ray topograph. To compare the structural patterns of a bio-molecule determined by AFM imaging and that derived from the 3D structure, we have constructed an *in silico* X-ray topograph using the crystal structure of TMV determined by X-ray fiber diffraction. We considered the distance from the solvent accessible surface of the molecule to the base plane as a height measure, thus the computed X-ray topograph represents such a quantity for TMV particles.[32] In brief, we placed a 3D atomic model of TMV in a rectangular grid. The atomic coordinates of a TMV subunit were obtained from the PDB code, 1VTM,[33] and those of a full helical repeat (49 subunits) were generated according to the transformation matrix therein. A 3D structure of three full helical repeats is presented in graphics (1) of Fig. 3a. For constructing the solvent accessible surface, we assigned the extended radius, *i.e.*, the sum of atomic van der Waals (vdw) and solvent radii, to each heavy atom. We used 1.5 Å for the solvent radius, and the atomic vdw radii were set to 2.18, 1.85, 1.8, 2.2 and 2.15 Å for C, N, O, S and P atoms,[34] respectively. To be comparable with the high-resolution AFM image acquired in this work, we set the pixel width of X-ray topograph (see image (2) of Fig. 3a) to 4.88 Å as well.

3 Results and discussion

In the present single-image analysis, geometric descriptions of the TMV particle were considered as coarse results from the raw AFM images, while determination of surface features is the ultimate goal. Quantification of the size of TMV will be performed on both AFM topographs with different probing resolutions by direct measurements of the intensities in the segmented region, and the surface features will be exposed by a series of image processing. Since the computed X-ray topograph represents the 2D surface information on the 3D detailed atomic coordinates of TMV, a comparison between AFM and X-ray techniques was made on their 2D topographs to reflect the 3D structural characteristics in the AFM image.

Descriptions of the outer appearance of TMV

A typical TMV particle possesses a body of rod shape about 300 nm long and 18 nm in diameter.[35] As revealed in EM micrographs, TMV exhibited a pattern of striation, reflecting the coat proteins arranged in a helical form.[28] One may see the striate pattern in the computed X-ray topograph in Fig. 3a as well.

An AFM topograph of TMV was generated by probing the surface of TMV coat proteins where mostly C-terminal and nearby amino acids are exposed. The AFM topographs of TMV are shown in images (1) of Fig. 1a and b for low and high resolutions, respectively. In the two AFM images, the striate pattern was barely discerned compared to the computed X-ray topograph. In each of Fig. 1a and b, images (2–5) correspond to the sequential outputs of computational procedures described in the Experimental section. To quantify TMV heights with a more accurate measurement, a plane fitting[36] was performed on each raw AFM image so that the tilts of the image base can be amended prior to data analysis, illustrated in image (2). Image (3) is a binary map composed of zero-crossing points which single out the TMV particle from the rest of image, and the segmented TMV is shown in image (4). In image (5), gray shades represent the TMV envelope. On top of the shading area the two eigenvectors of the intensity covariance matrix were plotted as well. They appear as two perpendicular white lines and are labeled as **u** and **v**.

The pixels in the **u** line are the reference data points for width estimations. Nevertheless, some width measurements were discarded due to either endpoint of the intersecting cross-segment connecting with the outside of the image, *e.g.*, the upper and lower parts of the TMV particle in the cropped part of the low-resolution AFM image. The width is obtained as the statistical average length of all possible vectors **v** (Fig. 1) of a given particle. The width was 30.96 ± 0.75 nm over 140 data points and 29.88 ± 0.07 nm over 510 data points from the low- and high-resolution images, respectively. Because AFM tips used in this study are not fully symmetric, it is possible that the difference in width is due to the scanning angle which is different in both images. Alternatively, the TMV size can also be represented by the virus height. From the low-resolution AFM image, the total height and the height of SAMs were calculated over 27 data points as 21.78 ± 0.08 nm and 1.74 ± 0.26 nm, respectively. The two height values were obtained as 19.63 ± 0.18 and 1.27 ± 0.23 nm over 319 data points from the high-resolution AFM image. Consequently, the TMV height was determined to be 20.04 ± 0.27 and 18.36 ± 0.29 nm with respect to the low and high resolutions in the present AFM imaging.

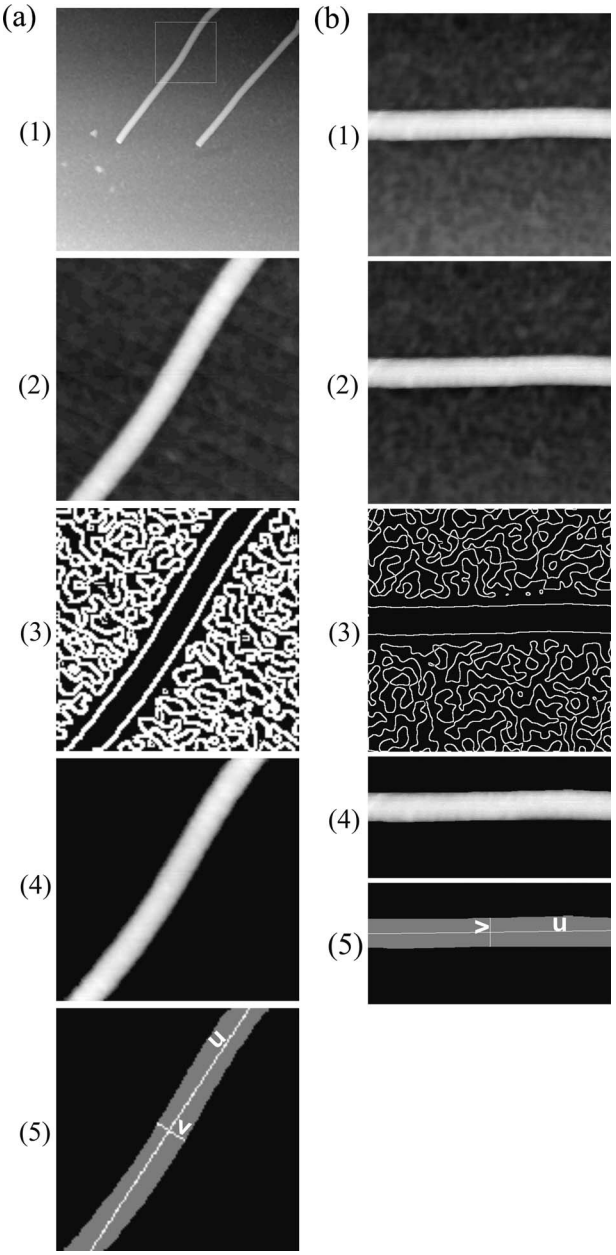

Fig. 1 (a) Intensity distribution of segmented TMV in the low-resolution AFM image. (1) AFM raw image of 512 × 512 pixels with a physical size of 1.0 × 1.0 μm^2, the intensity range is within [–33.8, 11.3] nm. (2) Sub-image of (1) with plane fitting: 128 × 128 pixels outlined by the white square whose top left corner locates at pixel (209, 30) of the AFM raw image with an intensity range of [–26.0, –3.8] nm. (3) Identified zero-crossing points from (2) with parameters $\Delta = 0.1$, $N_{mask} = 7$ and $\sigma = 3.5$. (4) Segmented TMV from (2). (5) Two characterizing vectors for TMV intensity distribution, **u** and **v**, see the Experimental section. (b) Intensity distribution of segmented TMV in the high-resolution AFM image. (1) AFM raw image of 512 × 512 pixels with a physical size of 0.25 × 0.25 μm^2, the intensity range is within [−6.5, 15.5] nm. (2) Plane fitting of the AFM raw image with the intensity range of [−4.6, 16.0] nm. (3) Identified zero-crossing points from (2) with parameters $\Delta = 0.001$, $N_{mask} = 7$ and $\sigma = 3.5$. (4) Segmented TMV from (2). (5) Two characterizing vectors for TMV intensity distribution, **u** and **v**, see the Experimental section.

As a reference, the TMV width and height from the computed X-ray topograph were also calculated. The computations yielded a value of 20.99 ± 1.50 nm over 361 data points for width and 19.02 ± 0.01 nm over 135 data points for height. Note that no supporting height (*i.e.*, SAMs) was present in the computed X-ray topograph. Comparing with the results from the computed X-ray topograph, the TMV width determined from AFM is consistently greater by about 10 nm whereas the height values are comparable. Such broadening effects on the lateral shape of the object in AFM imaging have been known as a consequence of tip convolutions and related to the tip shape in use.[37,38] On the other hand, the tip artifact is less severe in the vertical orientation of probed object, leading to a more accurate measurement in TMV heights.

To take into account the curved property of the TMV shape, width and height were also determined simultaneously by fitting a cross-section of the virus rod into an ellipse. After subtracting the previously estimated height of SAMs, the width- and height-like diameters were obtained to be 32.72 ± 1.06 and 19.76 ± 0.16 nm over 42 data points from the low-resolution AFM image. From the high-resolution AFM image, we obtained 29.92 ± 0.08 and 18.62 ± 0.06 nm over 454 data points for the two quantities. Similar calculations were performed on the computed X-ray topograph over 352 data points, yielding 18.62 ± 0.27 and 19.14 ± 0.11 nm for the two diameters of fitting ellipse. All these data values are summarized in Table 1. The width deviations between AFM and X-ray topographies were greater from the ellipse fitting than from direct measurements of width and height. From the computed X-ray topograph, the fitted cross-section of the TMV particle was nearly like a circle, while from both AFM images, it was a distorted circle, probably caused by adsorption forces of the substrate material or external forces of the AFM tip applied to the TMV surfaces. Clearly, the observed width of the TMV particle in both AFM images indicates that the tip radius was far from its nominal value (2 nm) and closer to 6 nm as established previously.[39]

Characterization of TMV structural features

On determination of surface features, we first applied supervised *DeStripe*[40] to reduce stripe noises in the low- and high-resolution AFM images. As observed in the denoised images, there were still no salient structural patterns discernible due to the too small intensity variation relative to the total intensity of the TMV particle. In order to raise the intensity contrast of TMV surfaces, we used the histogram equalizer to transform intensities of the AFM images, presented in images (3-low, 3-high) of Fig. 2. Now in the high-resolution AFM image, the enhanced intensity representation displays some features which somewhat resemble aligned disks, yet in the low-resolution AFM image the results still display less definite structures. To delineate the revealed structures, we have identified feature boundaries in TMV surfaces based on the intensity enhanced

Table 1 Geometric characterization of the TMV shape[a]

	AFM low resolution	AFM high resolution	X-ray
Width	31.0 ± 0.8	29.9 ± 0.1	21.0 ± 1.5
Height	20.0 ± 0.3	18.4 ± 0.3	19.0 ± 0.1
Horizontal diameter[b]	32.7 ± 1.1	29.9 ± 0.1	18.6 ± 0.3
Vertical diameter[b]	19.8 ± 0.2	18.6 ± 0.1	19.1 ± 0.1

[a] All the data values are in the unit of nm. [b] The results of ellipse fittings, see the Experimental section.

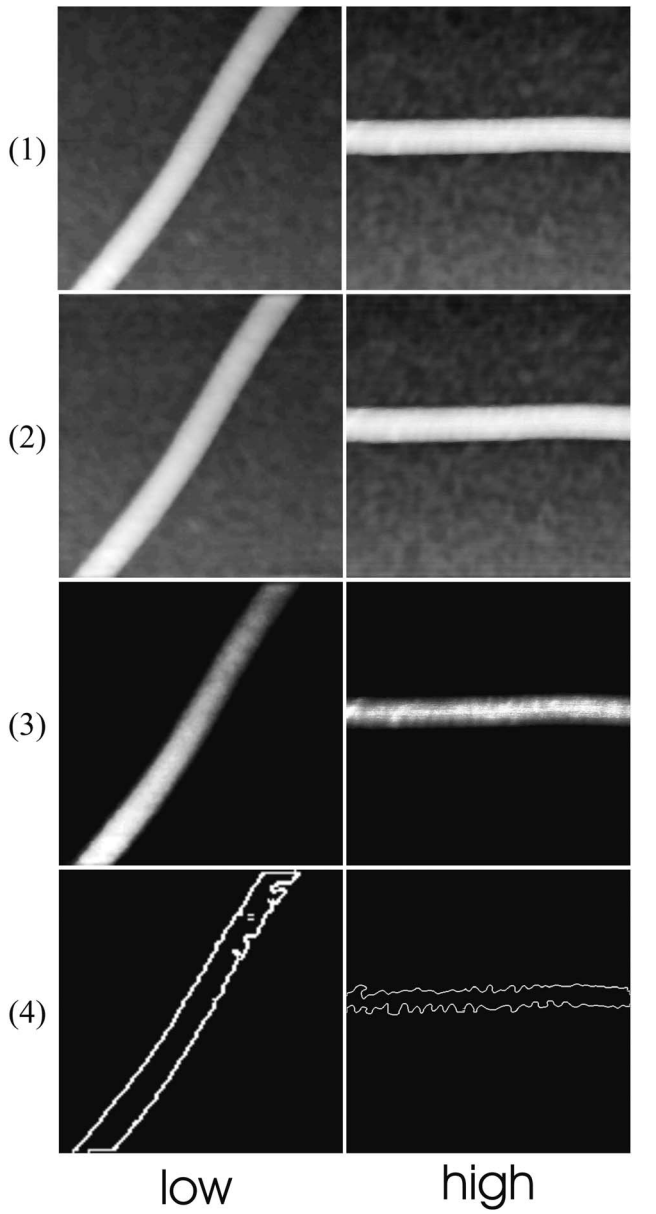

Fig. 2 Preliminary analysis on the low- and high-resolution AFM images described in Fig. 1. (1) Left: cropped low-resolution AFM image with the intensity range of [0.0, 28.6] nm. Right: the high-resolution AFM image with the intensity range of [0.0, 21.9] nm. (2) Left: denoised image for the low resolution with the intensity range of [−9.9, 19.9]. Right: denoised image for the high resolution with the intensity range of [−6.7, 15.3] nm. (3) Enhancement of intensity contrast in (2)-left and (2)-right by histogram equalization. The intensity values were normalized to [0.0, 1.0]. (4) Configurations of zero-crossing points for both AFM probing resolutions. The parameter values of the zero-crossing detector are $\Delta = 10^{-4}$, $N_{mask} = 3$ and $\sigma = 2.5$ for both images (3)-left and (3)-right.

AFM images. In accordance with image (3-low) of Fig. 2, there was no interpretable delineation of feature in the low-resolution image, see image (4-low) in Fig. 2. Not surprisingly, the boundary pixels found in image (4-high) of Fig. 2 from the high-resolution image outlined the disk-like structures already seen in image (3-high). These lines of evidence suggest that a further exaggeration of intensity gradation may disclose some obscure TMV surface features from AFM images.

Relationship of D_{xy}^2 features to probed heights

The second-order derivative operators are known to be more sensitive to intensity variations than the first-order. The D_{xy}^2 operation was tested first on the X-ray topograph of TMV to see if any finer detail can be revealed from the virus surfaces by this operation. In the crystal structure of TMV, the protein subunit, the single helical turn and the full helical repeat are three basic features of concern. Image (2) of Fig. 3a represents the TMV surfaces generated by the X-ray technique, which were computed according to a TMV 3D structure as partially illustrated in graphics (1) of Fig. 3a. From graphics (1), the separation between two adjacent helical turns was measured at about 2.5 nm. On a structural view of image (2), the maximum intensity extension of a helical turn is equivalent to the limit of one protein subunit, estimated to be ~2.96 nm. Across the surface of one protein subunit, the sizes of outmost protrusions ranged from 1.5 to 1.0 nm. Regarding the full helical repeat, the longitudinal length of three complete turns was estimated to be about 6.8–7.7 nm. The results of D_{xy}^2 operation on the

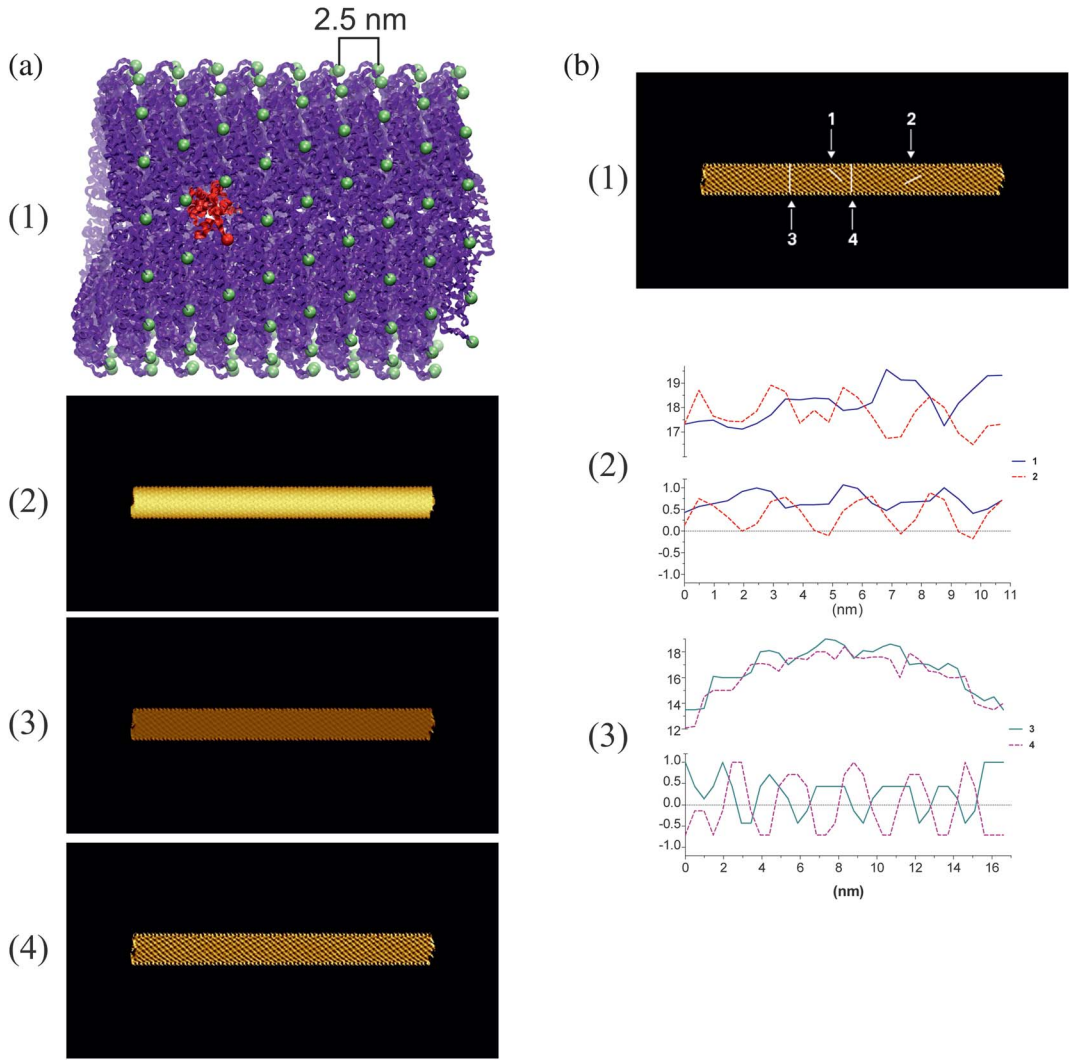

Fig. 3 (a) (1) Ribbon representation of the TMV crystal structure with six helical turns. One single subunit of the coat protein was highlighted in red. The C^{α} atoms of protein C-termini were drawn in green CPK. In the text, the distance between adjacent helical turns was measured between these C^{α} atoms. (2) *In silico* X-ray topography of TMV assembly with 63 helical turns of protein subunits. The computed topograph consists of 512 × 256 pixels with a pixel width of 4.88 Å and an intensity range of [0.0, 19.16] nm. (3) The D_{xy}^2 representation for the X-ray topography with the dimensionless intensity range of [−1.0, 1.0]. (4) Enhanced image of (3) by histogram equalization with the dimensionless intensity range of [−1.0, 1.0]. (b) Relationship of height and enhanced D_{xy}^2 values for the TMV particle. (1) Locations of compared segments in topographic and enhanced D_{xy}^2 images from the *in silico* X-ray topography of TMV, labeled as 1–4. (2) A plot of height and enhanced D_{xy}^2 values from selections 1 and 2. (3) A plot of height and enhanced D_{xy}^2 values from selections 3 and 4. In both plots (2 and 3), the two upper curves for height are in the unit of nm, while the two lower ones for enhanced D_{xy}^2 data are dimensionless.

computed X-ray topograph are presented in image (3) of Fig. 3a. Mathematically, the greater the D_{xy}^2 magnitude, the more dramatic the local intensity changes.

We further augmented the intensity contrast of the D_{xy}^2 image by histogram equalization. In image (4) of Fig. 3a, the enhanced D_{xy}^2 image exhibits a pattern like a bead net. Each bead coincides with a local maximum in the computed X-ray topograph, and reflects the location of C-terminus of a single coat protein. To relate enhanced D_{xy}^2 data to topographic measures, we have performed metric analysis on both image representations for feature determination. Four locations in the top panel of Fig. 3b were sampled to exemplify the compatibility of observations between enhanced D_{xy}^2 and topographic data computed from the X-ray technique. Both selections 1 and 2 passed through diagonal maxima of enhanced D_{xy}^2 values but oriented somewhat orthogonally. Selections 3 and 4 were placed crossing the virus rod. Right below image (1) of Fig. 3b, plot (2) presents height and enhanced D_{xy}^2 curves from selections 1–2; the two upper curves represent heights and the lower ones correspond to the enhanced D_{xy}^2 data. Note that the B-spline interpolation method in Gwyddion was adopted for plotting all illustrating selections throughout the paper. In general, the enhanced D_{xy}^2 curves display a saw-tooth trend where each tooth represents one structure and is well separated from two others beside it. The tooth size is about 2.5 nm on average, which matches with the C^{α}–C^{α} distance of two adjacent helical turns in the crystal structure of TMV. Since selections 3 and 4 cross the virus rod, their tooth numbers were found to be in agreement with the number of C-termini in one face of a TMV helical turn, colored in green in graphics (1) of Fig. 3a.

An inspection of the reciprocal relationship between height and enhanced D_{xy}^2 data shows that the two quantities from selection 1 are somewhat out of phase, while those from selection 2 are approximately in phase. In other words, a peak of the enhanced D_{xy}^2 curve locates in between two peaks of the height curve from selection 1, while from selection 2, the location of a peak in the enhanced D_{xy}^2 curve coincides approximately with that in the height curve. Similar tendencies were found in plot (3) for selections 3 and 4. These results show that the enhanced D_{xy}^2 operation is capable of not only delineating the feature border but also highlighting the most prominent part of the object surface. Although one may distinguish the surface structure from the height curve, the metric size obtained from the enhanced D_{xy}^2 curve is more regular and thus easier to be quantified. For example, around 4.5 nm in the spanning coordinate of selection 2, a small bulge existing in the height curve disappeared from the corresponding enhanced D_{xy}^2 curve. To the human eye, the enhanced D_{xy}^2 image provides a neater representation than the topograph for analyzing structural patterns. The present results show that a combination of D_{xy}^2 operation and histogram equalization on a topographic image enables to amplify surface features of an imaged bio-particle.

Exploring fine details of the TMV surface in AFM images

From the previous experience in treating the computed X-ray topography, we then applied D_{xy}^2 operation and histogram equalization to the two denoised AFM images. The results are shown in Fig. 4a, where images (1 and 2) are the processed outcomes of the low-resolution image while (3 and 4) correspond to the high resolution image. Comparison of image (4-low) of Fig. 2 and image (2) of Fig. 4a from the low-resolution AFM image shows that no perceptible pattern was found in the former image while some structures resembling piled disks appeared in the latter. This indicates the importance of appropriateness in processing tools used for revelation of a structural pattern. Moreover, the same structural pattern can be perceived at different AFM probing resolutions in both image (2) of Fig. 4a and image (3-high) of Fig. 2.

Examining image (4) of Fig. 4a and image (3-high) of Fig. 2, we found that the enhanced D_{xy}^2 operation on the high-resolution AFM image led to finer scales of observable features than

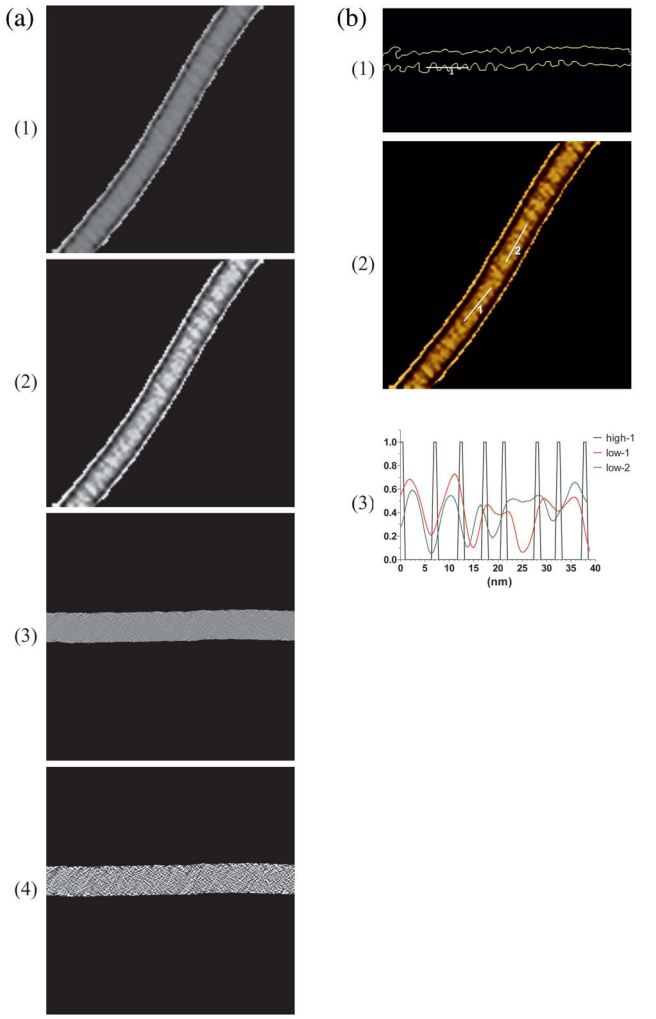

Fig. 4 (a) (1) The D_{xy}^2 image from the low-resolution AFM image. (2) The enhanced D_{xy}^2 image from the low-resolution AFM image. (3) The D_{xy}^2 image from the high-resolution AFM image. (4) The enhanced D_{xy}^2 image from the high-resolution AFM image. (b) Size comparison of structural patterns in the low- and high-resolution AFM images. (1) Location of selection high-1 indicated in the map of zero-crossing points from the high-resolution AFM image. (2) Locations of selections low-1 and low-2 specified in the enhanced D_{xy}^2 representation from the low-resolution AFM image. (3) A plot of dimensionless data values from selections high-1, low-1 and low-2.

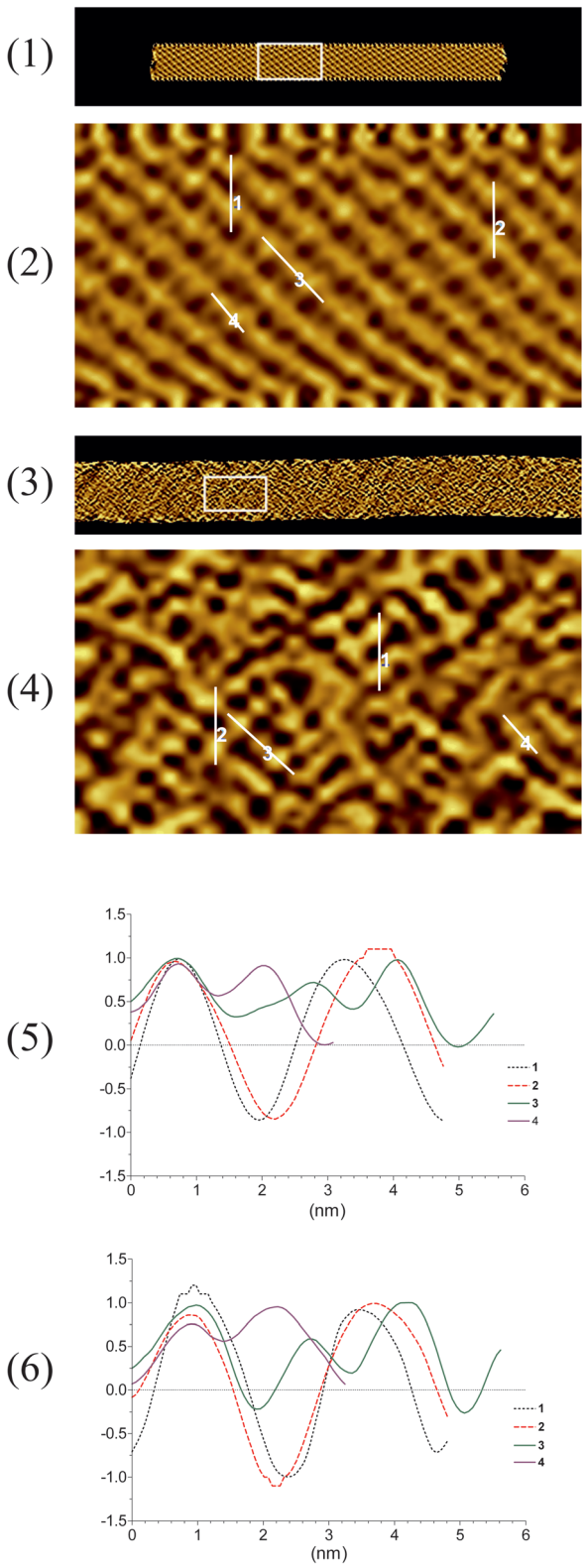

Fig. 5 Comparison of TMV surface features from X-ray and AFM topographies. (1) The location of the compared region, from the X-ray enhanced D_{xy}^2 image, is circled by a white rectangle. (2) The zoom-in of the compared region described in (1) with four selections 1–4. The zoomed image was magnified eight times of the white rectangle. (3) The compared region of the AFM enhanced D_{xy}^2 image is also outlined by a white rectangle. (4) The zoom-in of the white rectangle in image (3) with four selections 1–4. The zoom-in was also magnified eight times of the compared region. (5) A plot of data values from selections 1–4 of the X-ray enhanced D_{xy}^2 image. (6) A plot of data values from selections 1–4 of the AFM enhanced D_{xy}^2 image. Rendering performed with Gwyddion.[36]

the histogram equalization alone. This comparison demonstrates that with the same AFM image, different processing effects may reveal different scales of surface features. Comparison of either images (4-low, high) of Fig. 2 or images (2 and 4) of Fig. 4a shows that with the same processing tools the scale of the revealed feature correlates with the probing resolution used in AFM imaging. As expected, the higher the probing resolution, the finer the feature probed. In conclusion, the probing resolution and data processing are two crucial factors for revealing surface features of a bio-particle in AFM imaging.

We also quantified the size of disk-like patterns that appeared in the AFM images by different data processing and with two probing resolutions. Images (1 and 2) of Fig. 4b show the locations of sampled segments from the high- and low-resolution AFM images, respectively. In plot (3) of Fig. 4b, the enhanced D_{xy}^2 curves of selections low-1 and low-2 fit qualitatively with selection high-1. From metric analysis on selection high-1, the gap distance between two zero-crossing points was in the range of 3.8–6.8 nm over eight data points. The varying gaps reflect fluctuations in aligning helical turns of the TMV assembly. Among them, the longest distance was comparable to the length of one full helical repeat, whereas the rest infer a substructure to the full helical repeat.

Implications of 2D surface features in the 3D molecular structure

By way of its construction, the computed X-ray topography naturally contains the information of a detailed atomic model for the TMV particle. Taking advantage of the processing effects of the enhanced D_{xy}^2 operation, we compared the enhanced D_{xy}^2 representation, instead of topography, for the high-resolution AFM image with that for the X-ray topography. From images (1–4) of Fig. 5, both the enhanced D_{xy}^2 images for X-ray and AFM globally have a similar quality on the structural pattern, yet in detail the AFM structures are in less regular order. Recall that the computed X-ray topograph was made based on a rigid, perfect arrangement of TMV protein subunits, the AFM image, on the other hand, represents the topography of a real virus exposed to air at the time of data acquisition. This structural fluctuation is one major origin, besides any other instrumental problem, to make a direct match between X-ray and AFM techniques difficult.

To get around this problem, we sampled different regions in the two compared images (2 and 4) of Fig. 5 where local intensity distributions are analogous. In images (2 and 4), selections 1 and 2 orient vertically and link two granular structures with similar distance values, while selections 3 and 4 position diagonally and pass through three granular structures. From the enhanced D_{xy}^2 representation of the computed X-ray topography, the granular structure locates a protein subunit of TMV. Plots (5 and 6) of Fig. 5 present the data values from sampled regions for X-ray and AFM techniques, respectively.

The curves of selections 1 and 2 in both plots contain only two significant peaks due to the presence of two granular structures. Based on the peak-to-peak distance, the size of one granular structure from selection 1 was estimated to be 2.5 nm for both X-ray and AFM, while that from selection 2 is 3.1 nm for X-ray and 2.8 nm for AFM. Consequently, the size of the granular structure, which implicates a protein subunit, is about 2.5–2.8 nm by AFM, very close to what was found in the 3D crystal structure of TMV. Equivalently, selections 1 and 2 in image (4) may represent a partial helical turn probed by AFM. In terms of feature-based resolution,[12] the high-resolution AFM imaging was able to determine the most basic structural feature of TMV, *i.e.*, the protein subunit.

As derived from the computed X-ray topograph (*cf.* Fig. 3a), the corresponding enhanced D_{xy}^2 representation displays hereditary features of substructures in between adjacent protein subunits in TMV surfaces. Selections 3 and 4 in image (2) of Fig. 5 were sampled for representing such structural features finer than the protein subunit. We also searched in image (4) of Fig. 5 for comparable segments and labeled the same to link with image (2) for the AFM part. From selection 3, there are three peaks in the plots (5 and 6) and thus two peak-to-peak distances to be evaluated. We obtained (2.1, 1.28) nm for X-ray and (1.74, 1.47) nm for AFM. Although there are some discrepancies in the data values, the curve tendencies from X-ray and AFM are similar. From selection 4, only one peak-to-peak distance was of concern, and the value was the same for both X-ray and AFM, $\sim$1.3 nm. These results show that the scale of structural features extracted from the AFM image is up to 1.3 nm, reaching the high-resolution standard (<2 nm) in AFM imaging. More important, the study implies that the acquired AFM images do carry the 3D information of the TMV structure and this information is deducible.

4 Conclusions

Our objective was to link the surface features to the 3D structure of a single virus from AFM imaging. The interactions of molecular subunits, which underlie the mechanism of formation of the virus assembly, can be interpreted from topographic measurements by AFM. Therefore, an image representation was sought to resolve the 2D structure of a virus on a nanoscale, next to the atomic scale. We have investigated probing resolutions and processing algorithms which enable us to reveal detailed features of the TMV surface from AFM height measurements. As a coarse result, shape descriptions of the virus particle are usually the main presentation for an imaging study, while detailed surface features are often overlooked due to non-triviality of image visualization. The present work shows that the fine features in TMV surfaces are extractable by performing AFM imaging at a high probing resolution and processing the image with enhanced D_{xy}^2 operation.

Since the TMV shape has been well characterized by various techniques, it may be used for estimating tip effects on AFM-determined surfaces of bio-particles. As illustrated in the enhanced D_{xy}^2 representation for the high-resolution AFM image, structural patterns clearly cannot be manifested by multi-image averaging owing to pattern irregularity and deviation in each single image, despite the highly ordered TMV in the crystal structure. The single-image analysis provides an alternative to extract valuable information from a biological system suffering from structural flexibility during the imaging procedure. This strategy aims to loosen strictness on sample conditions in AFM for resolving structural features at high resolutions. At least, it should be applicable to a probed specimen similar to TMV, or it ought to be effective in molecular surface recognition[41] for studying a biological activity involving a complexed structure of multiple components.[42]

Acknowledgements

This work was partially supported by the French ANR agency (Grant ANR-P2N-2010-NANO-003) and by the French Alternative Energies and Atomic Energy Commission.

References

1 Z. Shao, J. Mou, D. M. Czajkowsky, J. Yang and J.-Y. Yuan, *Adv. Phys.*, 1996, **45**, 1–86.
2 P. Parot, Y. F. Dufrene, P. Hinterdorfer, C. Le Grimellec, D. Navajas, J.-L. Pellequer and S. Scheuring, *J. Mol. Recognit.*, 2007, **20**, 418–431.
3 I. Casuso, F. Rico and S. Scheuring, *J. Mol. Recognit.*, 2011, **24**, 406–413.
4 F. A. Schabert and A. Engel, *Biophys. J.*, 1994, **67**, 2394–2403.
5 J. M. Teulon, Y. Delcuze, M. Odorico, S.-w. W. Chen, P. Parot and J. L. Pellequer, *J. Mol. Recognit.*, 2011, **24**, 490–502.
6 *Dynamic Force Spectroscopy and Biomolecular Recognition*, ed. A. R. Bizzarri and S. Cannistraro, CRC Press, Boca Raton, 2012.
7 J. K. H. Hörber and M. J. Miles, *Science*, 2003, **302**, 1002–1005.
8 M.-H. Trinh, M. Odorico, M. E. Pique, J.-M. Teulon, V. A. Roberts, L. F. Ten Eyck, E. D. Getzoff, P. Parot, S.-w. W. Chen and J.-L. Pellequer, *Structure*, 2012, **20**, 113–120.
9 S. Scheuring, T. Boudier and J. N. Sturgis, *J. Struct. Biol.*, 2007, **159**, 268–276.
10 D. M. Czajkowsky and Z. Shao, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 14960–14965.
11 J. Frank, *Biopolymers*, 2003, **68**, 223–233.
12 P. Fechner, T. Boudier, S. Mangenot, S. Jaroslawski, J. N. Sturgis and S. Scheuring, *Biophys. J.*, 2009, **96**, 3822–3831.
13 J.-L. Pellequer and S.-w. W. Chen, *Proteins*, 2006, **65**, 192–202.
14 E. H. Egelman, *Ultramicroscopy*, 2000, **85**, 225–234.
15 H. Stahlberg, D. Fotiadis, S. Scheuring, H. Remigy, T. Braun, K. Mitsuoka, Y. Fujiyoshi and A. Engel, *FEBS Lett.*, 2001, **504**, 166–172.
16 S.-w. W. Chen and J. L. Pellequer, *BMC Struct. Biol.*, 2011, **11**, 7.
17 M. Husain, T. Boudier, P. Paul-Gilloteaux, I. Casuso and S. Scheuring, *J. Mol. Recognit.*, 2011, **25**, 292–298.

18 B. W. Erickson, S. Coquoz, J. D. Adams, D. J. Burns and G. E. Fantner, *Beilstein J. Nanotechnol.*, 2012, **3**, 747–758.
19 F. Schabert, A. Hefti, K. Goldie, A. Stemmer, A. Engel, E. Meyer, R. Overney and H.-J. Güntherodt, *Ultramicroscopy*, 1992, **42–44**, 1118–1124.
20 D. Fotiadis, *Curr. Opin. Biotechnol.*, 2012, **23**, 510–515.
21 S. Venkataraman, D. P. Allison, H. Qi, J. L. Morrell-Falvey, N. L. Kallewaard, J. E. Crowe, Jr and M. J. Doktycz, *Ultramicroscopy*, 2006, **106**, 829–837.
22 Y. F. Drygin, O. A. Bordunova, M. O. Gallyamov and I. V. Yaminsky, *FEBS Lett.*, 1998, **425**, 217–221.
23 F. Zenhausern, M. Adrian, R. Emch, M. Taborelli, M. Jobin and P. Descouts, *Ultramicroscopy*, 1992, **42–44**, 1168–1172.
24 G. Stubbs, S. Warren and K. Holmes, *Nature*, 1977, **267**, 216–221.
25 K. Namba and G. Stubbs, *Science*, 1986, **231**, 1401–1406.
26 C. Sachse, J. Z. Chen, P. D. Coureux, M. E. Stroupe, M. Fandrich and N. Grigorieff, *J. Mol. Biol.*, 2007, **371**, 812–835.
27 M. A. Ramin, G. Le Bourdon, K. Heuze, M. Degueil, C. Belin, T. Buffeteau, B. Bennetau and L. Vellutini, *Langmuir*, 2012, **28**, 17672–17680.
28 R. W. Horne, J. M. Hobart and R. Markham, *J. Gen. Virol.*, 1976, **31**, 265–269.
29 P. J. Butler, *J. Gen. Virol.*, 1984, **65**, 253–279.
30 D. Marr, *Vision*, W.H. Freeman & Co, New York, 1982.
31 R. C. Gonzalez and R. E. Woods, *Digital image processing*, Pearson Education, Inc., New Jersey, 2008.
32 S.-w. W. Chen and J. L. Pellequer, *Nucleic Acids Res.*, 2013, **41**, W412–W416.
33 R. Pattanayek and G. Stubbs, *J. Mol. Biol.*, 1992, **228**, 516–528.
34 S.-w. W. Chen, M. H. V. Van Regenmortel and J.-L. Pellequer, *Curr. Med. Chem.*, 2009, **16**, 953–964.
35 P. J. Butler, *Philos. Trans. R. Soc., B*, 1999, **354**, 537–550.
36 Gwyddion, http://gwyddion.net/, accessed 6 August 2013.
37 M. J. Allen, N. V. Hud, M. Balooch, R. J. Tench, W. J. Siekhaus and R. Balhorn, *Ultramicroscopy*, 1992, **42–44**, 1095–1100.
38 J. Yang, J. Mou, J. Y. Yuan and Z. Shao, *J. Microsc.*, 1996, **182**, 106–113.
39 M.-H. Trinh, M. Odorico, L. Bellanger, M. Jacquemond, P. Parot and J.-L. Pellequer, *J. Mol. Recognit.*, 2011, **24**, 503–510.
40 J. L. Pellequer, Y. J. Chen and S. w. W. Chen, *Removal of non-uniform stripe noises from AFM images*, Macau, China, 2012, DOI: 10.1109/iCBEB.2012.335.
41 E. Katchalski-Katzir, I. Shariv, M. Eisenstein, A. A. Friesem, C. Aflalo and I. A. Vakser, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 2195–2199.
42 J. L. Pellequer, A. J. Gale, E. D. Getzoff and J. H. Griffin, *Thromb. Haemostasis*, 2000, **84**, 849–857.

